

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2005, 00:48:39 ; Search time 254 Seconds
(without alignments)

8078.311 Million cell updates/sec

Title: US-09-600-714-41
Perfect score: 1254
Sequence: 1 atgagctctagaccgcgcg.....attgctgtgattttaa 1254

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUG_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1185.2	94.5	1384	1 US-08-553-888A-1	Sequence 1, Appl1
2	1182	94.3	1466	1 US-08-553-888A-2	Sequence 2, Appl1
3	177.8	14.2	1900	4 US-09-949-016-1457	Sequence 1457, Ap
4	175.4	14.0	1792	4 US-09-799-451-362	Sequence 362, App
5	165.4	13.2	1939	4 US-09-949-016-3778	Sequence 3778, Ap
6	109.6	8.7	1650	4 US-09-461-325-75	Sequence 75, Appl
7	109.6	8.7	1650	4 US-10-012-542-75	Sequence 75, Appl
8	109.6	8.7	1650	4 US-10-115-123-75	Sequence 75, Appl
9	56.6	4.5	34230	4 US-09-949-016-12052	Sequence 12052, A
10	56.6	4.5	128470	4 US-09-949-016-13765	Sequence 13765, A
11	56.2	4.5	601	4 US-09-949-016-49987	Sequence 49987, A
12	56.2	4.5	35676	4 US-09-949-016-13199	Sequence 13199, A
13	53.8	4.3	29172	4 US-09-949-016-15520	Sequence 15520, A
14	52	4.1	601	4 US-09-949-016-49982	Sequence 49982, A
15	50.4	4.0	134008	4 US-09-949-016-13841	Sequence 13841, A
16	47.4	3.6	64610	4 US-09-949-016-12214	Sequence 12214, A
17	45.4	3.6	39154	4 US-09-949-016-12384	Sequence 12384, A
18	45.4	3.6	39154	4 US-09-949-016-12801	Sequence 12801, A
19	45.4	3.6	39443	4 US-09-949-016-14326	Sequence 14326, A
20	44.6	3.6	39443	4 US-09-949-016-14327	Sequence 14327, A
21	44.6	3.6	152132	4 US-09-949-016-13845	Sequence 13845, A
22	44.6	3.6	152145	4 US-09-949-016-12371	Sequence 12371, A
23	44.4	3.5	31391	4 US-09-949-016-14319	Sequence 14319, A
24	44	3.5	522	4 US-09-949-016-103758	Sequence 103758, A
25	44	3.5	30656	4 US-09-949-016-14613	Sequence 14613, A
26	42.8	3.4	10304	4 US-09-627-465B-1	Sequence 1, Appl1
27	42.8	3.4	254405	4 US-09-949-016-14381	Sequence 14381, A

28	41	3.3	484	4 US-09-621-976-14663	Sequence 14663, A
29	41	3.3	540	4 US-09-902-540-4263	Sequence 4263, Ap
30	41	3.3	2094	4 US-09-252-991A-16118	Sequence 16118, A
31	41	3.3	2616	4 US-09-252-991A-16486	Sequence 16486, A
32	41	3.3	6909	3 US-09-199-637A-111	Sequence 111, App
33	41	3.3	10644	4 US-09-902-540-1028	Sequence 1028, Ap
34	40.8	3.3	1218	4 US-09-902-540-5379	Sequence 5379, Ap
35	40.8	3.3	34552	4 US-09-902-540-1262	Sequence 1262, Ap
36	40.6	3.2	1422	4 US-09-252-991A-8351	Sequence 8351, Ap
37	40.6	3.2	1503	4 US-09-252-991A-8488	Sequence 8488, Ap
38	40.6	3.2	1761	4 US-09-252-991A-8423	Sequence 8423, Ap
39	40.6	3.2	142783	4 US-09-949-016-15127	Sequence 15127, A
40	40.2	3.2	729	4 US-09-489-039A-5837	Sequence 5837, Ap
41	40.2	3.2	1389	4 US-09-902-540-2697	Sequence 2697, Ap
42	40.2	3.2	16265	4 US-09-902-540-1126	Sequence 1126, Ap
43	40	3.2	801	4 US-09-252-991A-6902	Sequence 6902, Ap
44	40	3.2	1410	4 US-09-252-991A-6871	Sequence 6871, Ap
45	40	3.2	1482	4 US-09-252-991A-6850	Sequence 6850, Ap

ALIGNMENTS

```
RESULT 1
US-08-553-888A-1
; Sequence 1, Application US/0853888A
; Patent No. 5723293
; GENERAL INFORMATION:
; APPLICANT: Huang
; TITLE OF INVENTION: DIAGNOSTIC METHOD AND KIT FOR
; TITLE OF INVENTION: DETERMINING RH BLOOD GROUP GENOTYPE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553, 888A
; FILING DATE: 11/06/95
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Dea, Sean W.
; REGISTRATION NUMBER: 37690
; REFERENCE/DOCKET INFORMATION: 454-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1384 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-553-888A-1

Query Match 94.5%; Score 1185.2; DB 1; Length 1384;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 121; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1 ATGAGCTTAAGAACCCGCGATGTCGCGGCGTGCCTGCGGCGCTTAACACTG 60
DB 31 ATGAGCTTAAGAACCCGCGATGTCGCGGCGTGCCTGCGGCGCTTAACACTG 90
61 GAAGCAGCTTCATCTCTCTCTTATTTTAAACCACTAAGACGCTTCTTAAGAGAT 120
```

D	b		91	GAAAGAGTCTCATTCCTCCTCTTCATTTATTCCACATATGACGCTTCTTAGAAGAT	150
O	y		121	CAAAAAGGGCGTCGTGCGACTCTTAATCAAGTGCGCAAGATCTGAACCGTATGCGGCCATT	180
D	b		151	CAAAAAGGGCGTCGTGCGACTCTTAATCAAGTGCGGCCAAGATCGATGACCGTAGAGCGGCCTT	210
O	y		181	GCGTTGGGCTTCACACTCGAGTTTCCGAGACACAGCTGAGACATGTGGGCTTCAAC	240
D	b		211	GCGTTGGGCTTCCTCACTCAATTTCCGAGACACAGCTGAGACATGTGGGCTTCAAC	270
O	y		241	CTCTTCAATGCTGCGGCTTTGGTGTGCAATGCGCAATCTGTGGAACGGCTTCTGAAGCAG	300
D	b		271	CTCTTCAATGCTGCGGCTTTGGTGTGCAATGCGCAATCTGTGGAACGGCTTCTGAAGCAG	330
O	y		301	TTCCTCTCTGGGAAGTGATCATCACACTGTTCAGTATTTGGGCTGGGACAATGAGTGTCT	360
D	b		331	TTCCTCTCTGGGAAGTGATCATCACACTGTTCAGTATTTGGGCTGGGACAATGAGTGTCT	390
O	y		361	TGTGTGGTGTGATCTGAGTGAATGCTGTCTTGGGGAAGGTCACTTGGCGAGTTGGTG	420
D	b		391	ATGTGCGTGTGATCTGAGGCGGCTGTCTTGGGGAAGGTCACTTGGCGAGTTGGTG	450
O	y		421	GTGATGGTGTGATGAGGAGTGAAGAAGTTTAAAGCACTTAAGATGTGATCAAGTAATATC	480
D	b		451	GTGATGGTGTGATGAGGAGTGAAGAAGTTTAAAGCACTTAAGATGTGATCAAGTAATATC	510
O	y		481	TTCAACACAGACTTACCATGAACATGATGCATCTACGTGTCCAGCGTATTTGGG	540
D	b		511	TTCAACACAGACTTACCATGAACATGATGCATCTACGTGTGTCCAGCGTATTTGGG	570
O	y		541	CTGTCTGTGGCCGTGGGCGCTGCCAAGGCTCTTACCAGAGGAACGAGAGATTAATATCAG	600
D	b		571	CTGTCTGTGGCCGTGGGCGCTGCCAAGGCTCTTACCAGAGGAACGAGAGATTAATATCAG	630
O	y		601	AAGCAACGATNCCAGTTGTGTGCGATGCTGGGCGCCCTCTTCTGTGAGATGTTCTGG	660
D	b		631	AAGCAACGATNCCAGTTGTGTGCGATGCTGGGCGCCCTCTTCTGTGAGATGTTCTGG	690
O	y		661	CCAAGTTCACTCTGCTCTGCTGAGAGTCCAATGAAAGAAAGATCCGTGTTCAAC	720
D	b		691	CCAAGTTCACTCTGCTCTGCTGAGAGTCCAATGAAAGAAAGATCCGTGTTCAAC	750
O	y		721	ACCTACTATGCTGATGACATGACGCGTGGAGACGACATCCAGGGATCATCTTGGCTCAC	780
D	b		751	ACCTACTATGCTGATGACATGACGCGTGGAGACGACATCCAGGGATCATCTTGGCTCAC	810
O	y		781	CCCCAAGGGAAGATGACAAAGCTTATGTGCAACAGTCCGCTGTTGGACAGAGCCGTGCT	840
D	b		811	CCCCAAGGGAAGATGACAAAGCTTATGTGCAACAGTCCGCTGTTGGACAGAGCCGTGCT	870
O	y		841	GTGGGATCCTCGTGCACCTGATCCCTTCTCCGTGGCTTGCATGCTGTGGCTTTGTG	900
D	b		871	GTGGGATCCTCGTGCACCTGATCCCTTCTCCGTGGCTTGCATGCTGTGGCTTTGTG	930
O	y		901	GCTGGGCTGATCTCCGTGGGGAAGCCAAATGACCTGCGGGGTGTTTGAACCGAGTGTG	960
D	b		931	GCTGGGCTGATCTCCGTGGGGAAGCCAAATGACCTGCGGGGTGTTTGAACCGAGTGTG	990
O	y		961	GGGATTTCCCACACAGTCCATATGAGGCTCAACATTCAGCTTGTGGGTCTCTTGGAGAG	1020
D	b		991	GGGATTTCCCACACAGTCCATATGAGGCTCAACATTCAGCTTGTGGGTCTCTTGGAGAG	1050
O	y		1021	ATCATCTACATTTGCTGCTGGTGTATATCCGTCCGAGCGCGCAATGGCATATTTGGC	1080
D	b		1051	ATCATCTACATTTGCTGCTGGTGTATATCCGTCCGAGCGCGCAATGGCATATTTGGC	1110
O	y		1081	TTCCAGGTCTCTCTAGCATTTGGGGAATCAGCTTGGCCATCTGATAGCTCTTCAAGTCT	1140
D	b		1111	TTCCAGGTCTCTCTAGCATTTGGGGAATCAGCTTGGCCATCTGATAGCTCTTCAAGTCT	1170
O	y		1141	GGTCTCTGACAGGTTTGTCTCTTAATTTTAAAAATAAGAAAGAACCTCATGTGGGCTAAA	1200
D	b		1171	GGTCTCTGACAGGTTTGTCTCTTAATTTTAAAAATAAGAAAGAACCTCATGTGGGCTAAA	1230

QY	1201	TATTGGATGACCAAGTTTCTTGAGAGTTTCCCATTTGGCCCTGGATTTTAA	1254
Db	1231	TATTGGATGACCAAGTTTCTTGAGAGTTTCCCATTTGGCCCTGGATTTTAA	1284
		RESULT 2	
		US-08-553-888A-2	
		/ Sequence 2, Application US/0855388A	
		/ Patent No. 5723293	
		/ GENERAL INFORMATION:	
		/ APPLICANT: Huang	
		/ TITLE OF INVENTION: DIAGNOSTIC METHOD AND KIT FOR	
		/ DETERMINING RH BLOOD GROUP GENOTYPE	
		/ NUMBER OF SEQUENCES: 3	
		/ CORRESPONDENCE ADDRESS:	
		/ ADDRESSEE: Hoffmann & Baron, LLP	
		/ STREET: 350 Jericho Turnpike	
		/ CITY:	
		/ STATE: New York	
		/ COUNTRY: USA	
		/ ZIP: 11753	
		/ COMPUTER READABLE FORM:	
		/ MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage	
		/ COMPUTER: IBM compatible	
		/ OPERATING SYSTEM: MS-DOS	
		/ SOFTWARE: wordperfect	
		/ CURRENT APPLICATION DATA:	
		/ APPLICATION NUMBER: US/08/553, 888A	
		/ FILING DATE: 11/06/95	
		/ CLASSIFICATION: 435	
		/ ATTORNEY/AGENT INFORMATION:	
		/ NAME: O'Dea, Sean W.	
		/ REGISTRATION NUMBER: 37690	
		/ REFERENCE/DOCKET NUMBER: 454-5	
		/ TELECOMMUNICATION INFORMATION:	
		/ TELEPHONE: (516) 822-3550	
		/ TELEFAX: (516) 822-3582	
		/ INFORMATION FOR SEQ ID NO: 2:	
		/ SEQUENCE CHARACTERISTICS:	
		/ LENGTH: 1466 nucleotides	
		/ TYPE: nucleic acid	
		/ STRANDEDNESS: double	
		/ TOPOLOGY: linear	
		US-08-553-888A-2	
		Query Match	
		Beet Local Similarity 94.3%; Score 1182; DB 1; Length 1466;	
		Matches 1209; Conservative 0; Mismatches 45; Indels 0; Gaps 0;	
QY	1	ATGAGCTTAAAGTACCCGCGGTGTCGCGGCGTGCCTTGGAGCCCTTAACATG	60
Db	45	ATGAGCTTAAAGTACCCGCGGTGTCGCGGCGTGCCTTGGAGCCCTTAACATG	104
QY	61	GAAGCAGTCTCATTTCTCTTTCTATTTTAAACCACTATGACGCTCTTAAAGAT	120
Db	105	GAAGCAGTCTCATTTCTCTTTCTATTTTAAACCACTATGACGCTCTTAAAGAT	164
QY	121	CAAAAGGGGCGTCGCGGATCCTATCAAGTTGGCCAAATCGACCGGATGGGCGCAT	180
Db	165	CAAAAGGGGCGTCGCGGATCCTATCAAGTTGGGCGCAATCTGACCGTGAATGGGCGCAT	224
QY	181	GGCTTGGGCTTCTCATCTCGAGTTTCCGAGACACAGCTGAGACAGTGTGGCCTTCAAC	240
Db	225	GGCTTGGGCTTCTCTCACTCAAAATTTCCGAGACACAGCTGAGACAGTGTGGCCTTCAAC	284
QY	241	CTCTTAAAGCGGCGTCTGATGACAGTGGGCAATCTGATGACCGCTTCCAGTACAG	300
Db	285	CTCTTAAAGCGGCGTCTGATGACAGTGGGCAATCTGATGACCGCTTCCAGTACAG	344
QY	301	TTCCTTCTTGGGAAGTGTATCACACTGTTCAATTCAGTATTCGCTGGGCAACATGAGTGT	360
Db	345	TTCCTTCTTGGGAAGTGTATCACACTGTTCAATTCAGTATTCGCTGGGCAACATGAGTGT	404

QY	361	TTGTCGGGTCGATCTCAGTGGATCTGCTCTGGGGAAAGTCAACTTGGCGAGTTGGG	420
Db	405	ATGTCGGTCGATCTCAACCGGCTCTGCTTGGGGAAAGTCAACTTGGCGAGTTGGG	464
QY	421	GTGATGTGTGCTGGTGGAGGTGACAGCTTTAAGCAACTGAGGATGTGTCATGTAATATC	480
Db	465	GTGATGTGTGCTGGTGGAGGTGACACTTTAAGCAACTGAGGATGTGTCATGTAATATC	524
QY	481	TTCAACAACAGACTACCAATGAACATGATGCACTTACGTGTTCGACACTTATTTTGGG	540
Db	525	TTCAACAACAGACTACCAATGAACCTGAGGCACTTACGTGTTCGACACTTATTTTGGG	584
QY	541	CTGTCGTGGCCCTGGTGCCTGGCAAAAGCCTCTAACCCGAGGAAACGAGAGATAAAGATCAG	600
Db	585	CTGACTGTGGCCCTGGTGCCTGGCAAAAGCCTCTAACCCGAGGAAACGAGAGATAATGATCAG	644
QY	601	ACAGCAACGATACCCAGTTTGTCTGCCATGCTGGGGCGCCCTCTTCTGTGATGTTCTGG	660
Db	645	AGAGCAACGATACCCAGTTTGTCTGCCATGCTGGGGCGCCCTCTTCTGTGATGTTCTGG	704
QY	661	CCAAGTTTCAACTGCTCTCTGCTGTAAGAAAGTCCAAATCGAAAGAAATCCCGTGTCAAC	720
Db	705	CCAAGTTTCAACTCTCTCTGCTGTAAGAAATCCAAATCGAAAGAAATGCAATGTTTCAAC	764
QY	721	ACCTACTATGCTGTAGCAGTCAGCCTGGTGACAGCCATCTCAGAGGTATCTTGGCTCAC	780
Db	765	ACCTACTATGCTCTAGCAGTCAGTGTGTGTGTGACAGCCATCTCAGAGGTATCTTGGCTCAC	824
QY	781	CCCCAAGGAAAGATCAGCAAGACTTATGTGCAACAGTGCAGTGTGTGACAGAGGCGTGCT	840
Db	825	CCCCAAGGAAAGATCAGCAAGACTTATGTGCAACAGTGCAGTGTGTGACAGAGGCGTGCT	884
QY	841	GTGGGTACTGTGTGACCTGATCCCTTCTCCGTGCTTGCATGTGTGCTGGGTCTTGTG	900
Db	885	GTGGGTACTGTGTGACCTGATCCCTTCTCCGTGCTTGCATGTGTGCTGGGTCTTGTG	944
QY	901	GCTGGGCTGATCTCCGTGGGGAGACCAAGTACTCCGCGGTGTGTAAACGAGTGCTG	960
Db	945	GCTGGGCTGATCTCCGTGGGGAGACCAAGTACTCCGCGGTGTGTGTAAACGAGTGCTG	1004
QY	961	GGGATTTCCCAACAGCTCATATGAGGCTCAAACTTCAAGCTTGTGTGGGTCTGCTGGAGAG	1020
Db	1005	GGGATTTCCCAACAGCTCATATGAGGCTCAAACTTCAAGCTTGTGTGGGTCTGCTGGAGAG	1064
QY	1021	ATCATCTATCATTTGTCGTGTGTGTGATATACGTCGGAACCGGCAATGGCAATGTGCG	1080
Db	1065	ATCATCTATCATTTGTCGTGTGTGTGATATACGTCGGAACCGGCAATGGCAATGTGCG	1124
QY	1081	TTCCAGTCTCTCTCAGCATTTGGGAAACTCAGCTTGGCCATGCTGTATGCTCAAGTCT	1140
Db	1125	TTCCAGTCTCTCTCAGCATTTGGGAAACTCAGCTTGGCCATGCTGTATGCTCAAGTCT	1184
QY	1141	GCTCTCTGACAGGTTTGCTCTTAATCTTAAATATGGAAGACCTCATGAGGCTAA	1200
Db	1185	GCTCTCTGACAGGTTTGCTCTTAATCTTAAATATGGAAGACCTCATGAGGCTAA	1244
QY	1201	TATTTGATGACCAAGTTTCTGGAAGTTTCCCATTTGGCTGTGGATTTTAA	1254
Db	1245	TATTTGATGACCAAGTTTCTGGAAGTTTCCCATTTGGCTGTGGATTTTAA	1298
RESULT 3			
US-09-949-016-1457			
Sequence 1457, Application US/09949016			
Patent No. 6812339			
GENERAL INFORMATION:			
APPLICANT: VENTER, J. Craig et al.			
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
FILE REFERENCE: CLO01107			
CURRENT APPLICATION NUMBER: US/09/949, 016			
CURRENT FILING DATE: 2000-04-14			

	PRIOR APPLICATION NUMBER: 60/241,755	
/	PRIOR FILING DATE: 2000-10-20	
/	PRIOR APPLICATION NUMBER: 60/237,768	
/	PRIOR FILING DATE: 2000-10-03	
/	PRIOR APPLICATION NUMBER: 60/231,498	
/	PRIOR FILING DATE: 2000-09-08	
/	NUMBER OF SEQ ID NOS: 207012	
/	SOFTWARE: FastSeq for Windows Version 4.0	
/	SEQ ID NO 1457	
/	LENGTH: 1900	
/	TYPE: DNA	
/	ORGANISM: Human	
US-09-949-016-1457		
Query Match	14.2%; Score 177.8; DB 4; Length 1900;	
Best Local Similarity	49.2%; Pred. No. 2,1e-43;	
Matches 544; Conservative	0; Mismatches 537; Indels 24; Gaps 2;	
Qy	131 TCCTGGCATCCCTATCAGATTGGGCCAATGCAACCGTGAATGGGGCCATTGGGCTTGAGCT	190
Db	167 TCTTTGGATTATACCCTGTTCCTGCATGATGTAATATTGTTGGGTTTGCT	226
Qy	191 TCCTCACCTGAGTTTCCGAGACACAGCTGAGACGTGGCCCTTCAACTCTTCAATGC	250
Db	227 TCCTCAGACCTTCTCTAGAATAATGGCTTCAGACATGGTGGATCACTACTCGTGG	286
Qy	251 TGCGCTTGGTGTGCAGTGGGCAATCCTGCTGACGGCTTCAGACCAAGTTCCTTCTG	310
Db	287 CTGCTTTGGGCTCTCAATGGGGGCACTATTGTACAGGAAATCTTG---CAAAGCCAGGGAC	343
Qy	311 GGAGGTGTATCACACTGTTCAGTATTCGGCTGGCCACATAGAGTCTTTGTGGGTC	370
Db	344 AGAATTTAACATTGGATCAAAAACATGATAATGACAGATTGATGTCAGCCACAGTTC	403
Qy	371 TGATCTCAGTGSATGCTGTCTTGGGGGAAGTCACTTGGCCGACAGTGGTGTGATGTGC	430
Db	404 TGAATCTTTTGGAGCTGTCTGGGAAAAACGAGCCCCAACCAATGCTATCATGACA	463
Qy	431 TGTGTAGGTGACAGCTTTTAGCCAAGCTGAGGATGTATCATGATATCTTTCAACAG	490
Db	464 TTTTAGAAATTTGTTTTCTTTGCCCAACATGATACTGTGTAGTGAATATTTTAAAGCCT	523
Qy	491 ACTTACCATGACATGATGATGACATCTACGTGTTCGACGCTATTTTGGGCTGTCTGG	550
Db	524 CTGACATTGGAGCATCATGACGATTCATGCTTTTGGGGCCTACTTTGGCTGTGAG	583
Qy	551 CCGTGGCTGGCCAAAGCCTCTACCCGAGGGAAACGAGATTAAGATCAGACAGACAGA	610
Db	584 CAGGACATCTTGATTCATCTTGGACTGGAAGAAAGGGGCAATGAATAAGAAAGTCCGATCT	643
Qy	611 TACCAGTTTGTCTGCATGTGTGGCCGCTCTTTCTTGTGGATTTCTGGSCAAGTTTCA	670
Db	644 ACTCAGACTGTGTGCAATGATGTGGAAGTCTTTCTGTGGATGTTTTGGCCAGCTTTA	703
Qy	671 ACTTCTCTCTGTGAGAAATCCAAATGAAAGAAATGCCGTGTTCACAACCTTACTATG	730
Db	704 ACTGGGCAATTTGCTGAACCTGAGACAAACAGTGCAGGGCCATTTGTAACACGTACTTCT	763
Qy	731 CTGTAGACAGTCAAGGTGTGACAGCCATCTCAGGGTATCTTGGGCTCACCCCAAGGGA	790
Db	764 CTCCTGCTGCTGTGTGTCTACAGCCTTTTGTCTTCTCAGCTCATGTGGAGACCGAGGCA	823
Qy	791 AGATCAGCAAGACTTATGTGTACAGTGCAGGTGTGGCAGAGAGCGTGGCTGTGGTACT	850
Db	824 AGCTCAACATGTTTCACTTCAGAAATGCACCTTGTGGAGGAGTGTGTGTGGGCACTT	883
Qy	851 CGTGTACCTGATCTCTTCTCGTGGCTTGCATGTGTGTGGGTCTTGTGTGGCTGTGA	910
Db	884 GTGGGATATGGAATTCACCACTTGTGTTCTATGATTAATGGAGCATTCAGAAATGG	943
Qy	911 TCTTCCGTGGGGGAGCAAGTACCTGGCGGGGATGTGAACCGAGTGTGGGATTTCCC	970
Db	944 TCTGTGTGCTTGGATACAGTTCTCTGACTCCACTTTTATCTACTTAACTGAGATTCATG	1003

QY	971	ACAGTTCATATGGGCTACAACTTCAGCTGCTGGGCTGGTGGAGATACATCTCA	1030
Db	1004	ATACTGTTGGGGTCCATACCTCCACGGCTTACCTGGTGTGTGGAGAGCCCTTGACGCA	1063
QY	1031	TTGTGCTGCTGCTGCTTATATACCGTCGAGGCCGGAGATGGCATATTTGGCTTCACAGTTC	1090
Db	1064	TTGTGGCAGTATGACATGGGCGCCCTTCAACAAGTCTATATGGCATGACAG-----	1111
QY	1091	TCCTCAGATATGGGGAACCTCAGCTTTGGGCATCGTATATGCTCTACGCTGCTGCTCTCA	1150
Db	1112	-----CAGCTGCACTGGGTTCTCTTATTCGGAACAGCAGTTGTTGGAAGTCTGAACTGA	1162
QY	1151	CAGGTTGCTCTTAATCTTAAAAATATGGAAGAACCCTCATGAGGCTTAAATATTTGATG	1210
Db	1163	CAGGTTTAAATCTTAAAGTTGCCTCTCTGGGAGACAGCATCTGACCAAGAACTGATATGATG	1222
QY	1211	ACCAAGTTTCTGGAAGTTGCTCTCA	1235
Db	1223	ATTCTGTTTATTTGGAAGGTCCCTTAA	1247

```

RESULT 4
US-09-799-451-362
: Sequence 362, Application US/09799451
: Patent No. 6783969
:
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Zhou, Ping
: APPLICANT: Goodrich, Ryle
: APPLICANT: Asundi, Vinod
: APPLICANT: Ren, Feiyun
: APPLICANT: Zhang, Jie
: APPLICANT: Xue, Aйдong J.
: APPLICANT: Zhao, Qing A.
: APPLICANT: Wang, Jіan-Rui
: APPLICANT: Ma, Yundong
: APPLICANT: Yamazaki, Victoria
: APPLICANT: Chen, Rui-hong
: APPLICANT: Wang, Zhiwei
: APPLICANT: Wang, Dunrui
: APPLICANT: Yang, Yonghong
: APPLICANT: Wehrman, Tom
: APPLICANT: Ghosh, Reena
: APPLICANT: Drmanac, Radcoje T.
:
: TITLE OF INVENTION: No. 6783969el Nucleic Acids and
: FILE OF INVENTION: Polypeptides
:
: FILE REFERENCE: 803
:
: CURRENT APPLICATION NUMBER: US/09/799,451
: CURRENT FILING DATE: 2001-03-05
: NUMBER OF SEQ ID NOS: 948
: SOFTWARE: pt_FL_genes Version 2.0
:
: SEQ ID NO 362
:
: LENGTH: 1792
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: NAME/KEY: CDS
:
: LOCATION: (32) .. (1354)
:
: US-09-799-451-362

```

Query Match	14.0%	Score 175.4	DB 4	Length 1792
Best Local Similarity	49.4%	Pred. No. 1.1e-42		
Matches 455	Conservative 0	Mismatches 466	Indels 0	Gaps 0
QY	141	CTATCAATTTGGCCAGATGTGACCGGATGGCGGCATGTGGCTTCCATCCTC	200	
Db	211	CTACCCAAAGCTTTCAGACGTCGACATGCCATGCTTGTGAGGCTTTGACTTCTTAAGGT	270	
QY	201	GAGTTTCCGAGACACACGCTGAGACGATGTGGCTTAACTTTCATGCTGAGCCTTGG	260	
Db	271	CTTCTGTGAGGGTTTACGGCTTTCAGACACGAGGGCTTACCTTCTCTCTGACCGCCTTTC	330	

QY	261	TTGTGAGTGGGCAATCTCTGCTGGACGGGCTTCTGTAGCCAGTTCCCTTCTGGGAAAGGTGGT	320
Db	331	CCTGGAGTGTGTCCACACTGTGTCCAGGGCTTTTCCACTCTCTTCCACGGTGGCCACATTC	390
QY	321	CATCAACACTGTTCAGTATTTGGGCTGGCCACCATGAGTGTCTTGTGGTGTGATCTCAGT	380
Db	391	TGTTGGCGGTGGAGAGCATGATCATGTGACTTTTGTGGGGGGCGGTGCTCATCTCCTT	450
QY	381	GGATGCTGTCTTTGGGGGAAGTTCACTTTGGCCGAGTTGTGTGTGATGTGTCTGTGGAGGT	440
Db	451	TGTGTCCCTTCTGGGCAAGACCGGGCTTACCCAGCTGTGCTCATAGGCCCTCTGTGAGGT	510
QY	441	GACAGCTTTAAGCAACTGAGGATGTATCATGATATATCTTCAACACAGACTACACAT	500
Db	511	GGTGTGTTTGGCATTATGATGATTTGTGTCTCTTATCTCTGGGGGTGAGAGATGGCGG	570
QY	501	GAACTATATGACATCTTCGTGTTCGACGCTTATTTTGGGCTGTCTGTGGCTGTGCT	560
Db	571	AGGCTCCATGACTATCCACACTTTGTGTGCTTACTTCCGGGCTGTCTTCCCGGGTTC	630
QY	561	GCCAAAGCCTTACCCCGAGGGAGCGGAGATTAAGATAGACAGACAGATACCGACTT	620
Db	631	GTACAGGCCCCAGCTGGAGAAAGAGACACACCGCCAGGGCTTCGTACCATTCAGACT	690
QY	621	GTCTGCCATGCTGGGCGGCTCTTCTTGTGATGTTCTGGCCAAATTTCAACTCTGCTCT	680
Db	691	CTTGGCCATGATTTGGGACCATCTTCTGTGATCTTGTGGCTTACACTTCAATGTGTGACT	750
QY	681	GCTGAGAAAGTCCAAATCCAAAGAAAGATGCGGTGTTCAACACTTATCTATGTCTTAGCACT	740
Db	751	CACAGCGCTGGGGGCTGGGAGCATTCGAGCGGCTTCAACATCTACTCTCCGTGGTGC	810
QY	741	CAGGCTGGTGACAGCCATCTCAGGGTATCTCTGGGCTACCCGCCCAAGGAAAGATAGCAA	800
Db	811	CAGACACCTTGGGACCTTTTGCTTGTCTGACGCTCTGTAGGGGAAAGTGGGAGGCTTGACAT	870
QY	801	GACTTATGTGCACAGTGGCGGTGTTTGGCAGAGAGCGGTGGCTTGGTACCTCGTGTACCT	860
Db	871	GGTCCACATCCAAATGCAAGCGGTGGCTGGAGGGGTGTGTGTGGGGACTTCAGATGAAT	930
QY	861	GATCCCTTTCGCGGCTTGGCAGATGTGTGGGTTCCTTGGCTGGGCTGATCTCCGTGG	920
Db	931	GATCTGACACCTTTTGGGCTCTGGGACGTGGCTTCTTGTGCTGGGACTGTCTCCACGCT	990
QY	921	GGAGCCCAAGTACTCGCCGGGTTTGTAAACGAGTGTGGGAGATTTCCCAACAGCTTCAT	980
Db	991	GGGGTAAAGTTTTCACGCCCACTTGTGATTCAAAATTCAAAGTCCAGACACATGTGG	1050
QY	981	CATGGGCTACAACTTCAGCTTGTGGGTCTGTGTGAGAGATTCATTCATATGTGTGCT	1040
Db	1051	AGTCCACAACTTCATGGGATGTGCGGGGGTCTTGGGGGCTCTCTGGGGGTCTTGTGGC	1110
QY	1041	GGTGCCTGATACGTTGGAGC	1061
Db	1111	TGGACTTGGCCACCATGAGC	1131

```

RESULT 5
US-09-949-016-3778
; Sequence 3778, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```

PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FaSeq for Windows Version 4.0
SEQ ID NO 3778
LENGTH: 1939
TYPE: DNA
ORGANISM: Human
US-09-949-016-3778

Query Match 13.2%; Score 165.4; DB 4; Length 1939;
Best Local Similarity 49.1%; Pred. No. 1.3e-39;
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

141 CTATCAAGTGGCCAGATCTGACCGTGAAGGCGCCATTGGCTGGAGTCTTCTCACTC 200
201 CTACCCAGGCTTCAGAGACGTGACCGTGAATGCTTGTGGGCTTGGCTTCTCAAGAC 260
201 GAATTTCCGAGACACAGCTGAGAGAGTGGCCTTCAACCTTTCATGCTGGCGCTTG 260
261 TTTCCTGACAGCGCTACCGCTTACGCGCGTGGCGCTTCACTTCTGTTGGCAGCTTCG 320
261 TGTGAGTGGGCAATCCTGCTGACGCGCTTCTGAGCCATTCCTTCTGGAAAGTGT 320
321 CATCCAGTGGCGCTGCTCAAGAGGCGTGTCTCACTTCTTCAAGACCGTACATGT 380
321 CATCAACGTGTGATGCTGGCTGGCCACCATGAGTGTCTTGTGGTGTGATCTCAGT 380
381 CGTGGGCGTGGAGAACTCATCAACGCTGACTTCTGGTGGCGCTGTCTGCGTGGCTT 440
381 GATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGGTGTGATGCTGTGTGAGGT 440
441 TGGGGAGTTCTGGTAAAGTCAAGCCCATTCAGCTGCTCATGACTTCTTCCAGT 500
441 GAGAGCTTTAGGCACTGAGATGTCATCAGTAATATCTTCAACACACATACCAT 500
501 GACCTCTTGTGCTGAATGATTCATCTTCAACCTCTAAAGGTGAAGATGCAAG 560
501 GAACATGATGACATCTACGTGTCGAGCCTATTTTGGGCTGTCTGTGGCTGTGCT 560
561 AGGCTCATGATCAGATCAGACATTTGGCGCTACTTTGGGCTCAGAGTGAACCGGATCT 620
561 GCCAAGCTCTTACCCGAGGAAAGAGATTAAGATCAGACAGCAACGATCCAGTTT 620
621 CTACCGACGCACTAGAGCAGAGCAAGAGAGACAAATTTGTGTAACAGTGGACT 680
621 GTGTGCAATGCTGGGCGCTCTTGTGAGATTTGTGGCAAGTTTCAACTGTGTCT 680
681 CTTTGCATGATTTGGCACCTCTTCTGTGATGTACTGGCCAGCTTCAACTCAGCAT 740
681 GCTGAGAACTCAATCGAAGAAAGAAATGCGTGTTCACACCTACTATCTGTAGCAT 740
741 ATCTCAATGAGGGAACGACAGACGAGCGCCCATATCAACTACTGCTCTTGGCAGC 800
741 CAGCGTGTGACACCATCTCAGAGGTATCTTGGCTCACCCCAAGGAAGATCAGCA 800
801 CTGCGTGTCTTACCTGGTGGCAATATCAGTGCCTTGCACAAAGAGGCAAGCTGACAT 860
801 GATTTATGTGCAAGTGGCGTGTGGAGAGAGCGCTGTGGTGTGATCTCTGTCACT 860
861 GGTGCAATCAAAATGCAACGCTCGAGAGAGGAGGCGTGGTATCCCTGTGAT 920
861 GATCCCTTCCGCTGCTGCTGATGCTGTGGTGTGATGCTGTGGTGTGATCTCCGCTG 920
921 GATGCTCATGCTTACGATGCTCTCATCATGCTGTGTGTGGGATCATCTCCACCT 980
921 GGGTTTTGTATACCTGACCCCATTTCTGAGTCCCGGCTGCATATCAGAGACATATG 980
981 CAGGGGTAAACTTCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1035
1041 CATTAACAATCTGCAATGCAATCTCGGCAATCATTAGGCGGATGCTGGTGTGTGTGT 1095

RESULT 6

US-09-461-325-75
Sequence 75; Application US/09461325A

Patent No. 6475753

GENERAL INFORMATION:

APPLICANT: Ruben et al.

TITLE OF INVENTION: 94 Human Secreted Proteins

FILE REFERENCE: P2029P1

CURRENT APPLICATION NUMBER: US/09/461,325A

CURRENT FILING DATE: 1999-12-14

EARLIER APPLICATION NUMBER: PCT/US99/13418

EARLIER FILING DATE: 1999-06-15

EARLIER APPLICATION NUMBER: 60/089,507

EARLIER FILING DATE: 1998-06-16

EARLIER APPLICATION NUMBER: 60/089,508

EARLIER FILING DATE: 1998-06-16

EARLIER APPLICATION NUMBER: 60/089,509

EARLIER FILING DATE: 1998-06-16

EARLIER APPLICATION NUMBER: 60/089,510

EARLIER FILING DATE: 1998-06-16

EARLIER APPLICATION NUMBER: 60/090,112

EARLIER FILING DATE: 1998-06-22

EARLIER APPLICATION NUMBER: 60/090,113

EARLIER FILING DATE: 1998-06-22

NUMBER OF SEQ ID NOS: 532

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 75

LENGTH: 1650

TYPE: DNA

ORGANISM: Homo sapiens

US-09-461-325-75

Query Match 8.7%; Score 109.6; DB 4; Length 1650;
Best Local Similarity 48.7%; Pred. No. 1.2e-22;
Matches 327; Conservative 0; Mismatches 344; Indels 1; Gaps 1;

364 TCGGCTGATATCTCAGTGAATGCTGTCTTGGGAAAGTCACTTGGCGCAGTTGTGTG 423
33 TCTGTCTGCTGCTCTTGGGCGAGTCTGTGGTAAGTCAAGCCCATTCAGCTGTATC 92
424 ATGTGCTGTGTGAGGTGACAGCTTATAGGCACTGAGATGTCTATCAATATCTTC 483
93 ATGACTTTCTTCAAGTGAACCTCTTCTGCTGATGATGATTCATCTTCACTGCTCA 152
484 AACAGACTTACCAATGAACATGATGACATCTTACGTGTGCAAGCTTATTTGGCTG 543
153 AAGGTAAAGATGACAGAGGCTTCATGACCATTCACACATTTGGCGCTTCTTGGGCTC 212
544 TCTGTGCGCTGCTGCTGCAAGGCTTACCCGAGGAAAGGAGTAAAGATCAAGACA 603
213 ACAGTAAACCGGATCTCTTACAGCAACCACTTGAAGAGCAAGAGAGCAAGATTTCT 272
604 GCAACGATACCCAGTTGTGTCTGCAATGCTGGGCGCTCTTCTTGTGATGTTGTGCA 663
273 GTGTACAGTGGAGCTCTTGTGCAATGATGAGCAACCTCTTCTGTGATGATCTGGCCC 332
664 AGTTTCACTTGTCTGTCTGAGAGTCCAATGAAAGAAATGCGTGTTCACACC 723
333 AGCTTCAACTCAGACCTATCTTCAATGAGGAGACAGCAGACCGGCAATCAACACC 392
724 TACTATGCTGTGACATGACGCGTGTGAGACGACCATCTCAGGATCATCTTGGCTACCCC 783
393 TACTGCTCTTGGAGACCTGTGCTTACTCTGCTGTGAGCAATATC-CAGTCCCTGCAAG 451
784 CAGGAGAAATCAGCAAGACTTATGTGCAAGTGCAGTGTGGCAGAGGCGTGTGTG 843
452 AAGGGCAAGGTGAGATGAGCAATCCAGAAATGCCAGCGCTGCAAGAGGAGTGGCGGTG 511
844 GGTACTGTGTGTCACTGATCTCTTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 903
512 GGTACCGCTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 571

QY 904 GGGCTGATCTCCGCGGGAGCCAGATACCTGCGGGGTGTGTAAACGAGTCTGGGG 963
 DB 572 GGCAATCTTCACCTCGGGTTTGTATACCTGACCCCATTCCTGAGTCCCGGCTGCAC 631
 QY 964 ATTCCCAAGCTCCATCATAGGGCTACACTTACGCTTGGGTCTGCTTGGAGATC 1023
 DB 632 ATCCAGACACATGTGGCATTAAACAATCTGCATGCAATCTTGGCATCATAGGGGATC 691
 QY 1024 ATCTACATGTG 1035
 DB 692 GTGGTCTGTG 703

RESULT 7
 US-10-012-542-75
 ; Sequence 75, Application US/10012542
 ; Patent No. 6627741
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: 94 Human Secreted Proteins
 ; FILE REFERENCE: P2029P1
 ; CURRENT APPLICATION NUMBER: US/10/012,542
 ; CURRENT FILING DATE: 2001-12-12
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
 ; NUMBER OF SEQ ID NOS: 532
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 75
 ; LENGTH: 1650
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-012-542-75

Query Match 8.7%; Score 109.6; DB 4; Length 1650;
 Best Local Similarity 48.7%; Pred. No. 1.2e-22;
 Matches 327; Conservative 0; Mismatches 344; Indels 1; Gaps 1;

QY 364 TCGGTGCTGATCTCAAGTGAATGCTCTTGGGGAAGTCAATTGGCGCAGTTGGGTG 423
 DB 33 TCTGTCTGCGTGGCCTTTGGGGCAGTTCTGGTAAGTACGCCCATTCAGCTGCTC 92
 QY 424 ATGTGCTGTGAGAGTGAACGCTTTRAGGCAACCTGAGATGGTCACTAAATATCTTC 483
 DB 93 ATGACTTCTTCCAAAGTACCTCTTCGTGTAATGATGATTCATTCCTTAACTGCTA 152
 QY 484 AACACAGACTACCAATGAACATGATGACATCTAGCTGTGGAGCTATTTTGGGCTG 543
 DB 153 AAGGTAAAGATCAGAGAGCTCCATGACATCAACATTTGGCGCTTACCTTTGGGCTC 212
 QY 544 TCTGTGCTGTGCTGCTGCCAAAGCCTCTAACCCGAGGGAACGAGAGTAAGATCAGAA 603
 DB 213 ACAGTAGCCGGATCTCTACCGAACCTTAGAGAGAGCAAGGAGAGCAAGATTTCT 272
 QY 604 GCAACATATCCAGTTTGTCTGCCATGCTGGGCGCCTCTCTTCTTGGATGTTCTGGCCA 663
 DB 273 GTGTACAGTGGACCTCTTGGCATGATTTGGCAACCTCTCTTCTGGATGTACTGGCCC 332
 QY 664 AGTTCAACTCTCTCTGCTGAGAGTCAATGGAAGGAAGAAATGCCGTGTTCAACACC 723
 DB 333 AGCTTCACTCAACCATATCTTCAATGAGGAGACGCAACGAGCCGCGCATCAACACC 392

QY 724 TACTATGCTGATGACAGTACAGTGTGAGACCATCTCAAGGATCATCTTGGCTCACCCC 783
 DB 393 TACTGCTCTGGCAACCTCGGCTGTACTACCTGGGTGGCAATATC-CAGTGCCCTGCAAG 451
 QY 784 CAAGGAAGATCAGCAAGACTTATGTCAOAGTGCCTGTGGCAGAGGCGTGGTGTG 843
 DB 452 AAGGGCAACTGGACATGTGTGCACATCCAGAAATGCCACCTCGCAGAGGCGGCGTG 511
 QY 844 GGTACCTGCTGACCTGATCCCTTCTCCGTGGCTTGCATGGTGTGCTGTGGCT 903
 DB 512 GGTACCGCTGCTGAGATGATGCTCATGCTTACGGTGCCTTATCATCGCTTCTGTCG 571
 QY 904 GGGCTGATCTCCGCGGGAGCCAGATACCTGCGGGGTGTGTAAACGAGTCTGGGG 963
 DB 572 GGCAATCTTCACCTCGGGTTTGTATACCTGACCCCATTCCTGAGTCCCGGCTGCAC 631
 QY 964 ATTCCCAAGCTCCATCATAGGGCTACACTTACGCTTGGGTCTGCTTGGAGATC 1023
 DB 632 ATCCAGACACATGTGGCATTAAACAATCTGCATGCAATCTTGGCATCATAGGGGATC 691
 QY 1024 ATCTACATGTG 1035
 DB 692 GTGGTCTGTG 703

RESULT 8
 US-10-115-123-75
 ; Sequence 75, Application US/10115123
 ; Patent No. 6774216
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: 94 Human Secreted Proteins
 ; FILE REFERENCE: P202930AP1D2
 ; CURRENT APPLICATION NUMBER: US/10/115,123
 ; CURRENT FILING DATE: 2002-04-04
 ; PRIOR APPLICATION NUMBER: PCT/US99/13418
 ; PRIOR FILING DATE: 1999-06-15
 ; PRIOR APPLICATION NUMBER: 60/089,507
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089,508
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089,509
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089,510
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/090,112
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090,113
 ; PRIOR FILING DATE: 1998-06-22
 ; NUMBER OF SEQ ID NOS: 532
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 75
 ; LENGTH: 1650
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-115-123-75

Query Match 8.7%; Score 109.6; DB 4; Length 1650;
 Best Local Similarity 48.7%; Pred. No. 1.2e-22;
 Matches 327; Conservative 0; Mismatches 344; Indels 1; Gaps 1;

QY 364 TCGGTGCTGATCTCAAGTGAATGCTCTTGGGGAAGTCAATTGGCGCAGTTGGGTG 423
 DB 33 TCTGTCTGCTGCGCCTTTGGGGCAGTTCTGGTAAGTACGCCCATTCAGCTGCTC 92
 QY 424 ATGTGCTGTGAGAGTGAACGCTTTRAGGCAACCTGAGATGGTCACTAAATATCTTC 483
 DB 93 ATGACTTCTTCCAAAGTACCTCTTCGTGTAATGATGATTCATTCCTTAACTGCTA 152
 QY 484 AACACAGACTACCAATGAACATGATGACATCTAGCTGTGGAGCTATTTTGGGCTG 543
 DB 153 AAGGTAAAGATCAGAGAGCTCCATGACATCAACATTTGGCGCTTACCTTTGGGCTC 212

NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 49987
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-49987

Query Match 4.5%; Score 56.2; DB 4; Length 601;
Best Local Similarity 56.1%; Pred. No. 1.3e-06;
Matches 106; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 613 CCCAGTTGTCGCCATCTGGGCGCCCTCTTTTGATGTTCTGCGCAAGTTTCAAC 672
DB 601 CCCCTCTTTGATTCACATACAGGGGACTCTTTCTGTGATGTTTGGCCAGCTTTAAC 542
QY 673 TCTGCTGCTGAGAGTCCATCGAAGAAATGCGGTTCACACTACTATGCT 732
DB 541 TCGGCATTGCTGAACCTGAGACAAAGTCAGGGGCATTGTAACACGTACTTCTCT 482
QY 733 GTAGACAGTCAGCGTGTGACAGCCATCTCAGGTCATCTGGCTCACCCCAAGGAG 792
DB 481 CTCGCTGCTGTGTGCTCAGAGCCTTTGCTTCCAGCTTAGTGAGACCCGAGGCAAG 422
QY 793 ATCAGCAG 801
DB 421 CTCACACATG 413

RESULT 12

US-09-949-016-13199
Sequence 13199, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13199
LENGTH: 35676
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(35676)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13199

Query Match 4.5%; Score 56.2; DB 4; Length 35676;
Best Local Similarity 56.1%; Pred. No. 1.7e-05;
Matches 106; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 613 CCCAGTTGTCGCCATCTGGGCGCCCTCTTTTGATGTTCTGCGCAAGTTTCAAC 672
DB 23970 CCCCTCTTTGATTCACATACAGGGGACTCTTTCTGTGATGTTTGGCCAGCTTTAAC 24029
QY 673 TCTGCTGCTGAGAGTCCATCGAAGAAATGCGGTTCACACTACTATGCT 732
DB 24030 TCGGCATTGCTGAACCTGAGACAAAGTCAGGGGCATTGTAACACGTACTTCTCT 24089
QY 733 GTAGACAGTCAGCGTGTGACAGCCATCTCAGGTCATCTGGCTCACCCCAAGGAG 792
DB 24090 CTCGCTGCTGTGTGCTCAGAGCCTTTGCTTCCAGCTTAGTGAGACCCGAGGCAAG 24149

QY 793 ATCAGCAG 801
DB 24150 CTCACACATG 24158

RESULT 13

US-09-949-016-15520
Sequence 15520, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15520
LENGTH: 29172
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(29172)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15520

Query Match 4.3%; Score 53.8; DB 4; Length 29172;
Best Local Similarity 62.0%; Pred. No. 8e-05;
Matches 85; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 153 CCAAGATCTGACCGATGAGGCGGCATTTGGGCTTCTCAGCTGAGTTCCGAG 212
DB 11588 CAGAGCGTGCACGATGATGCTTGTGGGCTTGGCTTCTCAGATTTCTGACAGCG 11647
QY 213 ACACAGCTGAGACAGTGTGGCCTTCAACCTCTTCATGCTGGGCGTGTGACAGTGGGC 272
DB 11648 CTACGGCTTCAGCGCGCGGTGCTTCAACTCTGTGTGACAGCTTGGGATTCAGTGGGC 11707
QY 273 AATCTGCTGACGGCT 289
DB 11708 GCTGCTCATGACAGGCT 11724

RESULT 14

US-09-949-016-49982/c
Sequence 49982, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 49982
LENGTH: 601

This Page Blank (uspto)

Db	1	ATGAGCTTAGTACC	CGGCTCTGTC	CGGCGCTGCT	CGCCCTCTG	GGCCCTTA	CACTG	60
QY	61	GAAGAGCTCTCAT	CTCTCTCTTAT	TTTTTTTAA	CCCACTA	TGACGCTCTTA	GAGAT	120
Db	61	GAAGAGCTCTCAT	CTCTCTCTTAT	TTTTTTTAA	CCCACTA	TGACGCTCTTA	GAGAT	120
QY	121	CAAAAAGGGGCT	GTGGGCACT	CTATCAG	TTGGCCAA	ATCTGAC	CCGTGAT	180
Db	121	CAAAAAGGGGCT	GTGGGCACT	CTATCAG	TTGGCCAA	ATCTGAC	CCGTGAT	180
QY	181	GGCTTGGGCTT	CGTACCT	CGAGTTT	CCGAGACA	CAGCTG	AGCAGT	240
Db	181	GGCTTGGGCTT	CGTACCT	CGAGTTT	CCGAGACA	CAGCTG	AGCAGT	240
QY	241	CTCTTCATGCT	GGGCGTT	TGGTGTG	SCATGGGCA	ATCTGCTG	AGCGCTT	300
Db	241	CTCTTCATGCT	GGGCGTT	TGGTGTG	SCATGGGCA	ATCTGCTG	AGCGCTT	300
QY	301	TTCCCTTCT	GGGAAAGTG	ATCATCA	CTGTTCA	GTTCAGT	TCGGCGG	360
Db	301	TTCCCTTCT	GGGAAAGTG	ATCATCA	CTGTTCA	GTTCAGT	TCGGCGG	360
QY	361	TTTGGCGGTG	CTGATCT	CATGAGAT	GCTGTCTT	GGGAAAG	GTCACTT	420
Db	361	TTTGGCGGTG	CTGATCT	CATGAGAT	GCTGTCTT	GGGAAAG	GTCACTT	420
QY	421	GTATGTGTG	TGTGAGGTG	AGACAGCTT	TAGGCA	CTGAGAT	TGTCAT	480
Db	421	GTATGTGTG	TGTGAGGTG	AGACAGCTT	TAGGCA	CTGAGAT	TGTCAT	480
QY	481	TTTCAACACAG	ACTAACCAT	GATGAACAT	GATGACAT	TACGTGT	TTCGACCT	540
Db	481	TTTCAACACAG	ACTAACCAT	GATGAACAT	GATGACAT	TACGTGT	TTCGACCT	540
QY	541	CTGTCTGTG	GGCCGTG	AGCTG	CCAAAGCC	CTAACCCG	AGGAA	600
Db	541	CTGTCTGTG	GGCCGTG	AGCTG	CCAAAGCC	CTAACCCG	AGGAA	600
QY	601	ACAGCAACG	ATACCACT	TTGTCTG	GCATGT	GGGCGCCT	CTTCT	660
Db	601	ACAGCAACG	ATACCACT	TTGTCTG	GCATGT	GGGCGCCT	CTTCT	660
QY	661	CCAAGTTT	CACTGCTG	CTGAGAG	AGTCCCA	TGAAAGAA	AGATG	720
Db	661	CCAAGTTT	CACTGCTG	CTGAGAG	AGTCCCA	TGAAAGAA	AGATG	720
QY	721	ACCTTACT	ATGCTGT	AGCACT	AGCGTGT	GACAGCAT	CTCAG	780
Db	721	ACCTTACT	ATGCTGT	AGCACT	AGCGTGT	GACAGCAT	CTCAG	780
QY	781	CCCAAGGG	AAATAGCA	GAAGCTT	ATATGT	CACTGCG	GGTGT	840
Db	781	CCCAAGGG	AAATAGCA	GAAGCTT	ATATGT	CACTGCG	GGTGT	840
QY	841	GTGGGTACT	GTGTCA	CTGATCCCT	CTCCGTG	CTTGTG	CAATG	900
Db	841	GTGGGTACT	GTGTCA	CTGATCCCT	CTCCGTG	CTTGTG	CAATG	900
QY	901	GCTGGGCT	GATCTCCG	TGGGGAG	CCCAAT	CTGCCGGG	GTGTAA	960
Db	901	GCTGGGCT	GATCTCCG	TGGGGAG	CCCAAT	CTGCCGGG	GTGTAA	960
QY	961	GGGATTTCC	CAAGCTCAT	ATGAGG	GTCAAACTT	CACTGCTG	GGGTG	1020
Db	961	GGGATTTCC	CAAGCTCAT	ATGAGG	GTCAAACTT	CACTGCTG	GGGTG	1020
QY	1021	ATCATTTCA	ATTTGTGCT	GTGCTG	ATATCG	GTGGAAG	CCGGCA	1080
Db	1021	ATCATTTCA	ATTTGTGCT	GTGCTG	ATATCG	GTGGAAG	CCGGCA	1080
QY	1081	TTCCAGTCT	CTCCAG	CAATTGG	GAAGCTA	CAAGCTT	GGCA	1140
Db	1081	TTCCAGTCT	CTCCAG	CAATTGG	GAAGCTA	CAAGCTT	GGCA	1140

Db	1081	TTCCAGGTCCTCCTCAGACATTTGGGAACTCAGCTTGCCANTGTAATGCTTCACTCT	11440		
Qy	1141	GGTCTCCTGACAGGTTTGTCTCTTAATCTTAAATATGAAAGCACTCATGAGGCTAA	12000		
Db	1141	GGTCTCCTGACAGGTTTGTCTCTTAATCTTAAATATGAAAGCACTCATGAGGCTAA	12000		
Qy	1201	TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTTGGATTTTAA	1254		
Db	1201	TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTTGGATTTTAA	1254		
RESULT 2	BD124102	1254 bp	DNA	linear	PAT 18-SEP-2002
LOCUS	BD124102				
DEFINITION	Novel nucleic acid molecule correlating to Rhesus weak D phenotype.				
ACCESSION	BD124102				
VERSION	BD124102.1	GI:23219047			
KEYWORDS	JP 2002500884-A/41.				
SOURCE	unidentified				
ORGANISM	unidentified				
REFERENCE	1 (bases 1 to 1254)				
AUTHORS	Fregel,V.A. and Wagner,F.F.				
TITLE	Novel nucleic acid molecule correlating to Rhesus weak D phenotype				
JOURNAL	Patent: JP 2002500884-A 41 15-JAN-2002; DRK BLUTSPENDEDIENST BADEN WÜRTTEMBERG GGMIBH				
COMMENT	OS Unidentified PN JP 2002500884-A/41 PD 15-JAN-2002 PF 18-DEC-1998 JP 2000528671 PR 23-JAN-1998 EP 98101203.2 PI VILLY A FREGEL,FRANZ F WAGNER PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/10,PC C12P21/02,C12P21/08,C12Q1/02,C12Q1/68,G01N33/566,C12N15/00,PC C12M5/00 CC Strandedness: Single; CC Topology: Linear; CC Novel nucleic acid molecule correlating to Rhesus weak D phenotype FH Key Location/Qualifiers FT CDS 1 1254. Location/Qualifiers 1..1254 /organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32544"				
FEATURES	source				
ORIGIN					
Query Match	100.0%;	Score 1254;	DB 6;	Length 1254;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 1254;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	ATGAGCTCTAAGTACC	CGCGGATCTGTCCGGCGTGCCTGCCCTCGGCGCCTTAACATCG	60	
Db	1	ATGAGCTCTAAGTACC	CGCGGATCTGTCCGGCGTGCCTGCCCTCGGCGCCTTAACATCG	60	
Qy	61	GAAAGCAGCTCTCATTTCTCTCTTTATTTTAA	CCCATATGACGCTTCTTGAAGGAT	120	
Db	61	GAAAGCAGCTCTCATTTCTCTCTTTATTTTAA	CCCATATGACGCTTCTTGAAGGAT	120	
Qy	121	CAAAAAGGCGTGTGGCATCTATGAAGTGGCCAA	ATTGACCTGATAGCGGCACTT	180	
Db	121	CAAAAAGGCGTGTGGCATCTATGAAGTGGCCAA	ATTGACCTGATAGCGGCACTT	180	
Qy	181	GCGTTGGGCTTCTCACTCGATTCGAGACA	CAGACTGAGACAGTGGCTTCAAC	240	
Db	181	GCGTTGGGCTTCTCACTCGATTCGAGACA	CAGACTGAGACAGTGGCTTCAAC	240	
Qy	241	CTCTTCATGCTGCGCTTGGTGTGTGATAGTGG	CAATCTGCTGACGCGCTTCTTACCCAG	300	
Db	241	CTCTTCATGCTGCGCTTGGTGTGTGATAGTGG	CAATCTGCTGACGCGCTTCTTACCCAG	300	

QY 301 TTCCTTCTGGGAAGGTGGTCATCACATGTTTCAGTATTCGAGCTGGCCACATGAGTGTCT 360
 DB 301 TTCCTTCTGGGAAGGTGGTCATCACATGTTTCAGTATTCGAGCTGGCCACATGAGTGTCT 360
 QY 361 TTGTGGTGTGATCTCAAGATGCTGTCTTTGGGGAAGGTCAACTTGGCGCATTTGGTG 420
 DB 361 TTGTGGTGTGATCTCAAGATGCTGTCTTTGGGGAAGGTCAACTTGGCGCATTTGGTG 420
 QY 421 GGTGATGTGTGGAGGTGAGTGAAGTGTGAGTGAAGTGTGAGTGAAGTGTGAGTGAAGT 480
 DB 421 GGTGATGTGTGGAGGTGAGTGAAGTGTGAGTGAAGTGTGAGTGAAGTGTGAGTGAAGT 480
 QY 481 TTCACACAGACTACACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
 DB 481 TTCACACAGACTACACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
 QY 541 TTGTCTGTGGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 600
 DB 541 TTGTCTGTGGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 600
 QY 601 ACAGCAACGATACCAAGTTTGTGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 660
 DB 601 ACAGCAACGATACCAAGTTTGTGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 660
 QY 661 CCAAGTTTCACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 720
 DB 661 CCAAGTTTCACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 720
 QY 721 ACCTACTATGCTGTAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 780
 DB 721 ACCTACTATGCTGTAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 780
 QY 781 CCCCAAGGGAAGTACGAGAACTTATGTCACAGTGTGTCATGTCATGTCATGTCATGTC 840
 DB 781 CCCCAAGGGAAGTACGAGAACTTATGTCACAGTGTGTCATGTCATGTCATGTCATGTC 840
 QY 841 GTGGGTAACCTGTGTCACTGTGTCACTGTGTCACTGTGTCACTGTGTCACTGTGTCACT 900
 DB 841 GTGGGTAACCTGTGTCACTGTGTCACTGTGTCACTGTGTCACTGTGTCACTGTGTCACT 900
 QY 901 GGTGGGCTGATCTCCGTGGGGAAGCCAGTACCTGCGGGGTGTGTAACCAAGTGTCTG 960
 DB 901 GGTGGGCTGATCTCCGTGGGGAAGCCAGTACCTGCGGGGTGTGTAACCAAGTGTCTG 960
 QY 961 GGGATTTCCCAAGCTTCATCATGAGGTGATGAGTGTGATGAGTGTGATGAGTGTGAT 1020
 DB 961 GGGATTTCCCAAGCTTCATCATGAGGTGATGAGTGTGATGAGTGTGATGAGTGTGAT 1020
 QY 1021 ATCATCTACATTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1080
 DB 1021 ATCATCTACATTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1080
 QY 1081 TTCAGGCTCTCTCTCAGACTTGGGGAAGTCTGAGTGTGCTGATGATGATGATGATG 1140
 DB 1081 TTCAGGCTCTCTCTCAGACTTGGGGAAGTCTGAGTGTGCTGATGATGATGATGATG 1140
 QY 1141 GGTCT 1200
 DB 1141 GGTCT 1200
 QY 1201 TATTTTGTATGACCAAGTTTCTGGAAGTTCCTCATTTGGCTGTGATTTTAA 1254
 DB 1201 TATTTTGTATGACCAAGTTTCTGGAAGTTCCTCATTTGGCTGTGATTTTAA 1254

RESULT 3
 AB018969 1254 bp mRNA linear PRI 03-JUL-1999
 LOCUS AB018969 Homo sapiens RndyO mRNA for Rh blood group D antigen (RHD),
 DEFINITION complete cds.
 ACCESSION AB018969
 VERSION AB018969.1 GI:5360898

KEYWORDS RndyO; Rh blood group D antigen (RHD).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 Hyodo, H., Ishikawa, Y., Kashiwase, K., Ogawa, A., Watanabe, Y., Tsuneyama, H., Toyoda, C., Uchikawa, M., Akaza, T. and Fujii, T.
 Polymorphisms of Rhdya in Japanese
 Unpublished
 2 (bases 1 to 1254)
 Uchikawa, M., Hyodo, H. and Ishikawa, Y.
 Direct Submision
 Submitted (20-OCT-1998) Hironobu Hyodo, Japanese Red Cross Central Blood Center, Research; 4-1-31, Hiroo, Shibuya, Tokyo 150-0012, Japan (E-mail:hyodo@hla.cb.c.jrc.or.jp, Tel: +81-3-5485-6009, Fax: +81-3-3406-7892)
 FEATURES
 Location/Qualifiers
 1..1254
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /isolate="YO"
 /db_xref="taxon:9606"
 1..1254
 /gene="RhdYO"
 /gene="RhdYO"
 /note="Rhd variant RndyO"
 /product="Rh blood group D antigen (RHD)"
 /protein_id="BA82159.1"
 /db_xref="GI:5360899"
 /creation="MSSKPRSVRCLPMLATLBAALLLPFFTHYDASLEDOGLVAVOYGDLTWAAIGLGLTSPFRHSMSVAFULPMLALGVMAILLDFLSOPSGKVVTLTFSILATMSALSVLSVDAVGLKXNLAIVLVVAVLALGVNLSINISKRVTHMMHIVPAHFGLSVYAVKLPRLPBGTDKXOTATPISLMSLGLPLMFWPENSALRLSPDKRNAVFNVTYVAVSVVVAISGLAHQKXISKTYVSAVLA GGVAVGTSHLIPSPMLAMVLGIVAGLISVGAKYLPGCCNRYLGIPIHSIMQYNSL LGIVAGIIVLVLDVTGAGNGMIGFYUJLSIGLSLAIIVIALTSGLLTGLLNLKI WKAPHBAKFFDDQVFWKPFHLAVGF"

ORIGIN

Query Match 99.9%; Score 1252.4; DB 9; Length 1254;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGCTTAAAGTACCCCGGCTGTGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
 DB 1 ATGAGCTTAAAGTACCCCGGCTGTGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
 QY 61 GAAGCAGCTCTAATCTCCCTCTTCTTATTTTAAACCACTAGAGCGCTTCTTAGAGAT 120
 DB 61 GAAGCAGCTCTAATCTCCCTCTTCTTATTTTAAACCACTAGAGCGCTTCTTAGAGAT 120
 QY 121 CAAGAAGGCTCTGAGCATCTTCAAGTGTGCAAGATCTGACCGTATGAGCGGCATT 180
 DB 121 CAAGAAGGCTCTGAGCATCTTCAAGTGTGCAAGATCTGACCGTATGAGCGGCATT 180
 QY 181 GGTGGGCTTCTCTACCTGAGTTTCCGAGACACAGCTGAGACAGTGTGGCTTTAAC 240
 DB 181 GGTGGGCTTCTCTACCTGAGTTTCCGAGACACAGCTGAGACAGTGTGGCTTTAAC 240
 QY 241 CTCTTATGCTGGCGCTGTGGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAG 300
 DB 241 CTCTTATGCTGGCGCTGTGGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAG 300
 QY 301 TTCCTTCTGGGAAGGTGGTCATCACATGTTTCAGTATTCGAGCTGGCCACATGAGTGTCT 360
 DB 301 TTCCTTCTGGGAAGGTGGTCATCACATGTTTCAGTATTCGAGCTGGCCACATGAGTGTCT 360
 QY 361 TTGTGGTGTGATCTCAAGATGCTGTCTTTGGGGAAGGTCAACTTGGCGCATTTGGTG 420
 DB 361 TTGTGGTGTGATCTCAAGATGCTGTCTTTGGGGAAGGTCAACTTGGCGCATTTGGTG 420

QY	421	GTGATGCTGCTGCTGAGGTGACAGCTTAAAGCACTGAGAGTGTCACTAATATC	480
Db	421	GTGATGCTGCTGCTGAGGTGACAGCTTAAAGCACTGAGAGTGTCACTAATATC	480
QY	481	TTTAAACAGACTACCAACATGAAATGATGACATTTACGTGTGACAGCTTATTTGG	540
Db	481	TTTAAACAGACTACCAACATGAAATGATGACATTTACGTGTGACAGCTTATTTGG	540
QY	541	CTGTCTGTGAGCTGTGCTGCTGCAAGAGCTTACCCGAGAGGAAACGAGATTAATAC	600
Db	541	CTGTCTGTGAGCTGTGCTGCTGCAAGAGCTTACCCGAGAGGAAACGAGATTAATAC	600
QY	601	ACAGCAGACATACCAAGTTTGTCTGACATGCTGAGGAGCCCTCTTGTGATGTTCTG	660
Db	601	ACAGCAGACATACCAAGTTTGTCTGACATGCTGAGGAGCCCTCTTGTGATGTTCTG	660
QY	661	CCAAGTTTCACTGCTGCTGCTGAGAGTCCATTCGAAAGAGAGATCCGTGTTCAAC	720
Db	661	CCAAGTTTCACTGCTGCTGCTGAGAGTCCATTCGAAAGAGAGATCCGTGTTCAAC	720
QY	721	ACCTAATATGCTGTAGACAGTACGCTGTGACAGCATCTCAGGATCTCTTGGCTCAC	780
Db	721	ACCTAATATGCTGTAGACAGTACGCTGTGACAGCATCTCAGGATCTCTTGGCTCAC	780
QY	781	CCCAAGGAGAGATGACAGAACTTATGTGACAGTGTGCTGTTGACAGAGCGTGCT	840
Db	781	CCCAAGGAGAGATGACAGAACTTATGTGACAGTGTGCTGTTGACAGAGCGTGCT	840
QY	841	GTGAGTACCTGCTGTGACAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	900
Db	841	GTGAGTACCTGCTGTGACAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	900
QY	901	GCTGAGCTGATCTCGTGTGAGGAGAGCAAGTACCTGCTGCTGCTGCTGCTGCTG	960
Db	901	GCTGAGCTGATCTCGTGTGAGGAGAGCAAGTACCTGCTGCTGCTGCTGCTGCTG	960
QY	961	GAGATTTCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1020
Db	961	GAGATTTCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1020
QY	1021	ATCATCTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1080
Db	1021	ATCATCTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1080
QY	1081	TTTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1140
Db	1081	TTTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1140
QY	1141	GCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1200
Db	1141	GCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1200
QY	1201	TATTTTGTATGACCAAGTTTGTGAGAGTTTCTCATTTGGCTGTGATTTTAA	1254
Db	1201	TATTTTGTATGACCAAGTTTGTGAGAGTTTCTCATTTGGCTGTGATTTTAA	1254

RESULT 4
HUMRHADANTI
LOCUS Human Rhd blood group antigen mRNA, complete cds. PRI 24-AUG-1993
DEFINITION L08429.1 GI:337390
VERSION 1
KEYWORDS Rhd blood group; Rhd blood group; antigen; blood group antigen.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1354)
Arce, M.A., Thompson, E.S., Wagner, S., Coyne, K.E., Ferdman, B.A. and
Lublin, D.M.
Molecular cloning of Rhd cDNA derived from a gene present in

QY	1	ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGCTGCTGCTGCTGCTGCTGCTGCTG	60
Db	1	ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGCTGCTGCTGCTGCTGCTGCTGCTG	60
QY	61	GAAGCAGCTCTCATTTCTCTCTTATTTTAAACCATATGACGCTTCTTGAAGAT	120
Db	61	GAAGCAGCTCTCATTTCTCTCTTATTTTAAACCATATGACGCTTCTTGAAGAT	120
QY	121	CAAAAGGCGCTGTGAGCATCTATCAAGTTGGCCAGATCTGACCTGATGCGCCATT	180
Db	121	CAAAAGGCGCTGTGAGCATCTATCAAGTTGGCCAGATCTGACCTGATGCGCCATT	180
QY	181	GAGTTGGGCTTCCACCTGAGTTTCCGAGACACAGCTGAGCAGTGTGCTTCAAC	240
Db	181	GAGTTGGGCTTCCACCTGAGTTTCCGAGACACAGCTGAGCAGTGTGCTTCAAC	240
QY	241	CTTTCATGCTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	300
Db	241	CTTTCATGCTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	300
QY	301	TTCCCTTCTGAGAGGTGTGCTATCACTGTTCAATGATTTGGCTGGCCACATGATCT	360
Db	301	TTCCCTTCTGAGAGGTGTGCTATCACTGTTCAATGATTTGGCTGGCCACATGATCT	360
QY	361	TTGTGCTGCTGATCTCAGTGTGATGCTGCTTGGGAGGTCACTTGGCGGAGTTGGG	420
Db	361	TTGTGCTGCTGATCTCAGTGTGATGCTGCTTGGGAGGTCACTTGGCGGAGTTGGG	420
QY	421	GTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	480
Db	421	GTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	480
QY	481	TTTAAACAGACTACCAACATGAAATGATGACATTTACGTGTGACAGCTTATTTGG	540
Db	481	TTTAAACAGACTACCAACATGAAATGATGACATTTACGTGTGACAGCTTATTTGG	540
QY	541	CTGTCTGTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	600
Db	541	CTGTCTGTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	600

ORIGIN

Query Match	99.9%	Score 1252.4	DB 9	Length 1354
Best Local Similarity	99.9%	Pred. No. 0		
Matches 1253	Conservative 0	Mismatches 1	Indels 0	Gaps 0

FEATURES

Source

1. 1354

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

1. 1254

/standard_name="Rhd"

/citation=(1)

/evidence=experimental

/product="Rhd blood group antigen"

/protein_id="AAA02679.1"

/db_xref="GI:337391"

/translation="MSKTPRSVRRLPLMALITLALILFFPTHYDASLEDOKGL
VASVQGDULTVMAALGLGLTSPRRHSWSVFNLMALGVOMAILDGLPSFP
SKRVITLFSRLATMSALSVLISVDVILGKVLALQVLYLVETALGNLRMTSNI
FNTDYHMMMHIVFAAYGLSVAMCPKPIPEGREDOTATTSIAMSICALPIWM
FWPSPSALRLSPRIERKNAVFNVTYAVVSVTALSGSLAHPOGSKISYVHSAVLA
GGVAVGSCHLIPSPMLAMVLTGVAGLISVGAATLPGCCNRRVGIPIHSSIMGVNPSL
LGLLEITTYIYILVITVGAAGNMGIFGVLLISBLISLIVIALTSLGLTLNLTKI
WAPRPAKFPDQVFWKPHLAVGP"

Qy	601	ACAGCAACGAATACCAAGTTTGTCTGACATGCTGGAGCGCCCTTCTTGTGGATGTCGG	660
Db	601	ACAGCAACGAATACCAAGTTTGTCTGACATGCTGGAGCGCCCTTCTTGTGGATGTCGG	660
Qy	661	CCAAGTTTCACCTGTGCTCTGCTGAGAAGTCCAAATCGAAGAAGAAATGCCGTGTTCAAC	720
Db	661	CCAAGTTTCACCTGTGCTCTGCTGAGAAGTCCAAATCGAAGAAGAAATGCCGTGTTCAAC	720
Qy	721	ACCTACTATGCTGTAGCAGTCAGCGTGTGACAGCCATCTCAGAGGTATCTTTGGCTTCAC	780
Db	721	ACCTACTATGCTGTAGCAGTCAGCGTGTGACAGCCATCTCAGAGGTATCTTTGGCTTCAC	780
Qy	781	CCCCAAGGGAAAGATCAGCAACAATTATGTGACAGTGGCGGTGTGGACGAGAGGGTGGCT	840
Db	781	CCCCAAGGGAAAGATCAGCAACAATTATGTGACAGTGGCGGTGTGGACGAGAGGGTGGCT	840
Qy	841	GTGGGTACCTGCTGTGTCACCTGATCCCTTCTCCGTGGCTTGCATAGTGTCGTGGCTTGTG	900
Db	841	GTGGGTACCTGCTGTGTCACCTGATCCCTTCTCCGTGGCTTGCATAGTGTCGTGGCTTGTG	900
Qy	901	GCTGGGCTGATCTCCGTGCGGGGAGCCAAAGTACCTGCCGGGTGTGTAAACCAAGTCTG	960
Db	901	GCTGGGCTGATCTCCGTGCGGGGAGCCAAAGTACCTGCCGGGTGTGTAAACCAAGTCTG	960
Qy	961	GGGATTTCCCAACAGCTCCATCATGAGGCTACAACATTCAGCTTGTGGCTTGGAGAG	1020
Db	961	GGGATTTCCCAACAGCTCCATCATGAGGCTACAACATTCAGCTTGTGGCTTGGAGAG	1020
Qy	1021	ATCATCTTACATATTGTGCTGCTGCTGCTGATACCGCTCGAGCGCGCAATGCGATATTGGC	1080
Db	1021	ATCATCTTACATATTGTGCTGCTGCTGCTGATACCGCTCGAGCGCGCAATGCGATATTGGC	1080
Qy	1081	TTCCAGAGTCTCTCTCAGCATTTGGGGAACTCAGCTTGGCCATCGTATAGCTTCAAGCTCT	1140
Db	1081	TTCCAGAGTCTCTCTCAGCATTTGGGGAACTCAGCTTGGCCATCGTATAGCTTCAAGCTCT	1140
Qy	1141	GGTCTCCCGAAGAGTTTGTCTCTCTTAATTTTAAATATGGAAAGCACTCATAGAGGCTTAA	1200
Db	1141	GGTCTCCCGAAGAGTTTGTCTCTCTTAATTTTAAATATGGAAAGCACTCATAGAGGCTTAA	1200
Qy	1201	TATTTTGTATGACCAAGTTTCTGGAAGTTTCCCTCATTTTGGCTGTGTGATTTTAA	1254
Db	1201	TATTTTGTATGACCAAGTTTCTGGAAGTTTCCCTCATTTTGGCTGTGTGATTTTAA	1254
RESULT 5			
LOCUS	HSRH11	1545 bp	mRNA linear PRI 07-JUL-1997
DEFINITION	H.sapiens mRNA for rhesus polypeptide (Rh11).		
ACCESSION	X63094.1	GI:36027	
VERSION	601994		
KEYWORDS	isoform; red cell membrane protein; Rh blood group; Rh polypeptide; Rhesus polypeptide.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 1545) Le Van Kim,C., Cherif-Zahar,B., Raynal,V., Mouru,I., Lopez,M., Cartion,J.P. and Colin.Y.		
TITLE	Multiple Rh messenger RNA isoforms are produced by alternative splicing		
JOURNAL	Blood 80		
MEDLINE	92360855		
PUBMED	1379850		
REFERENCE	2		
AUTHORS	Le van Kim,C., Mouru,I., Cherif-Zahar,B., Raynal,V., Cherrier,C., Cartion,J.P. and Colin.Y.		
TITLE	Molecular cloning and primary structure of the human blood group RhD polypeptide		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 89 (22), 10925-10929		
MEDLINE	93066356		
PUBMED	1438298		

REFERENCE 3 (bases 1 to 1545)
 AUTHORS Colijn, Y.
 TITLE Direct Submission
 JOURNAL Submitted (04-NOV-1991) Y. Colijn, INSERM, U76, INTS, 6, Rue
 A. Cabanef, 75015 Paris, FRANCE
 COMMENT See also X63094-98 & M44015.
 FEATURES
 source
 location/Qualifiers
 1..1545
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1p35"
 /clone="RH11"
 /tissue_type="bone marrow"
 /tissue_1b="human bone marrow cDNA"
 59..1312
 /codon_start=1
 /product="rhesus polypeptide RH1"
 /protein_id="CAA44808.1"
 /db_xref="GI:36028"
 /db_xref="GOA:Q02161"
 /db_xref="UniProt/Swiss-Prot:Q02161"
 /translation="MSKSPRSVRCLPMLTEALILFPEFTHYDASLPEDKGG
 VASVQODLTVMALIGLGLTSPSRHMSVAPNFMALGVOMALLDGLFSQPL
 SGKVTTLFSIRLATMSALSVLSVDYALGVKNLAQLVYMLVETRALIRMYISN
 FNTDYHMMWHIYVPAVFGLSVAMCKEPLPEGEDKQDATTIPSLAMGALFLM
 FWPENSNALRSPIERKNAVFNITYAAVSVTLISGSSLAPOQIKISIKYHSAVL
 GGVAVGTSCHLIPSPMLAMVLGVALISVGAKYLPCCCKNRVLGI PHSIINGYFSS
 LGLLEIYIVLIVLDTVAGNGMIGFQVLISIGLSIAIVIALTSGLLTGLLNLK
 WKAEHAKYFPDDQVFWKPHLAVP"
 ORIGIN
 Query Match 99.9%; Score 1252.4; DB 9; Length 1545;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0
 1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGGTGCTGCCCCCTTGCGCCTTAACACTG 60
 59 ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGGTGCTGCCCCCTTGCGCCTTAACACTG 118
 61 GAAGAGCTCATATTCCTCTCTCAATTTTTCACCACTAAGAGCTTCTTAAGAGAT 120
 119 GAAGAGCTCATATTCCTCTCTCAATTTTTCACCACTAAGAGCTTCTTAAGAGAT 178
 121 CAAAAGGGGCTCGTGGCATCTCATCAAGTTGGCCAAAGATCTGACCGTATGGCGGCATT 180
 179 CAAAAGGGGCTCGTGGCATCTCATCAAGTTGGCCAAAGATCTGACCGTATGGCGGCATT 238
 181 GCGTTGGGCTTCTTCACCTCGAGATTTCGGAGACACAGCTGAGAGAGTGTGGCTTCAAC 240
 239 GCGTTGGGCTTCTTCACCTCGAGATTTCGGAGAGACACAGCTGAGAGAGTGTGGCTTCAAC 298
 241 CTCTTCATGCTGGGCTGTGTGTGAGAGGGCAATCCGTCGAGAGGCTTCTCTGAGCCAG 300
 299 CTCTTCATGCTGGGCTGTGTGTGAGAGGGCAATCCGTCGAGAGGCTTCTCTGAGCCAG 358
 301 TTCCTTCTGGGAAGTGTGATCACTGTTCAAGTATTCGCTGGCCACCATGAGTGCT 360
 359 TTCCTTCTGGGAAGTGTGATCACTGTTCAAGTATTCGCTGGCCACCATGAGTGCT 418
 361 TTGTGCGTGTGATCTCAGTGGAGTGTCTTTGGGGAAAGTCACTTGGCGCAGTTTGCTG 420
 419 TTGTGCGTGTGATCTCAGTGGAGTGTCTTTGGGGAAAGTCACTTGGCGCAGTTTGCTG 478
 421 GTGATGTGCTGTGTGAGAGTGAACGCTTTAGGCAACCTGAGAGTGTATCACTGAATATC 480
 479 GTGATGTGCTGTGTGAGAGTGAACGCTTTAGGCAACCTGAGAGTGTATCACTGAATATC 538
 481 TTCAACACAGACTACCATGAACATGATGACATCTACGTTTGGCAGCCTAATTTTGGG 540
 539 TTCAACACAGACTACCATGAACATGATGACATCTACGTTTGGCAGCCTAATTTTGGG 598


```

Qy 541 CTGTCTGTGGCCCTGTGCTGCTGCCAAGACCTTACCCGAGGAAACGAGGATAAAGATCAG 600
Db 599 CTGTCTGTGGCCCTGTGCTGCTGCCAAGACCTTACCCGAGGAAACGAGGATAAAGATCAG 658
Qy 601 ACAGGACAGATACCCGAGGTTTGTCTGTGCAATGCTGGGCGCCCTCTTCTTGTGATGTTCTGG 660
Db 659 ACAGGACAGATACCCGAGGTTTGTCTGTGCAATGCTGGGCGCCCTCTTCTTGTGATGTTCTGG 718
Qy 661 CCAAGTTTCAACTCTGCTGCTGCTGAGAACTCCAAATGCAAGAAAGAAAGACCGCTGTTCAAC 720
Db 719 CCAAGTTTCAACTCTGCTGCTGCTGAGAACTCCAAATGCAAGAAAGAAAGACCGCTGTTCAAC 778
Qy 721 ACTTACTATGCTGTAGCACTGACCGTGTGACAGCAATCTCAGGCTCATCCTTGGCTCAC 780
Db 779 ACTTACTATGCTGTAGCACTGACCGTGTGACAGCAATCTCAGGCTCATCCTTGGCTCAC 838
Qy 781 CCCCAAGGGAATGACAGCAAGCTTATGTCACATGCGGCTTGGCAGAGGCGTGCT 840
Db 839 CCCCAAGGGAATGACAGCAAGCTTATGTCACATGCGGCTTGGCAGAGGCGTGCT 898
Qy 841 GTGGGTACCTGTGTACCTGATCCCTTCTCCGTGGCTTGCATGCTGCTGGTCTTGTG 900
Db 899 GTGGGTACCTGTGTGTACCTGATCCCTTCTCCGTGGCTTGCATGCTGCTGGTCTTGTG 958
Qy 901 GCTGGGCTGATCTCCGTGGGAGGACCAAGTACCTGCGGCGTGTGTAAACCAAGTCTG 960
Db 959 GCTGGGCTGATCTCCGTGGGAGGACCAAGTACCTGCGGCGTGTGTAAACCAAGTCTG 1018
Qy 961 GGAATTTCCCAACAGCTCCATGATGGGCTCAACACTTCACTGCGGCTTGGCTGAGAG 1020
Db 1019 GGAATTTCCCAACAGCTCCATGATGGGCTCAACACTTCACTGCGGCTTGGCTGAGAG 1078
Qy 1021 ATCATCTAATGTTGTGCTGCTGCTGCTGATACCGTGGAGCCGCAATGAGCATGATGGC 1080
Db 1079 ATCATCTAATGTTGTGCTGCTGCTGCTGATACCGTGGAGCCGCAATGAGCATGATGGC 1138
Qy 1081 TTCAGGCTCCCTCAGATGAGGAACTCAGCTTGGCCATCTGATAGCTTCAAGCTCT 1140
Db 1139 TTCAGGCTCCCTCAGATGAGGAACTCAGCTTGGCCATCTGATAGCTTCAAGCTCT 1198
Qy 1141 GGTCTCTGACAGGTTTGTCTTAAATCTTAAATATGAAAGCACTCTGAGGCTTAA 1200
Db 1199 GGTCTCTGACAGGTTTGTCTTAAATCTTAAATATGAAAGCACTCTGAGGCTTAA 1258
Qy 1201 TATTTATATACCAAGTTTCTGGAATTTCTTCATTTGCTGCTGATTTTAA 1254
Db 1259 TATTTATATACCAAGTTTCTGGAATTTCTTCATTTGCTGCTGATTTTAA 1312

```

```

RESULT 6
LOCUS HSRXIII 2790 bp mRNA linear PRI 07-JUL-1997
DEFINITION H.sapiens mRNA for rhesus polypeptide (RhxIII).
ACCESSION X63097.1 GI:36046
VERSION X63097.1 GI:36046
KEYWORDS isoform; red cell membrane protein; Rh blood group; Rh polypeptide;
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 2790)
Le Van Kim C., Cherif-Zahar B., Raynal V., Mouro I., Lopez M.,
Carton J.P. and Colin Y.
Multiple Rh messenger RNA isoforms are produced by alternative
splicing
Blood 80 (4), 1074-1078 (1992)
JOURNAL MEDLINE 92360855
PUBMED 1379850
REFERENCE 2
AUTHORS Le Van Kim C., Mouro I., Cherif-Zahar B., Raynal V., Cherrier C.,
Carton J.P. and Colin Y.
TITLE Molecular cloning and primary structure of the human blood group

```

```

RHD polypeptide
Proc. Natl. Acad. Sci. U.S.A. 89 (22), 10925-10929 (1992)
MEDLINE 9306356
PUBMED 1438298
REFERENCE 3 (bases 1 to 2790)
AUTHORS Colin Y.
TITLE Direct Submission
COMMENT Submitted (04-NOV-1991) Y. Colin, INSERM, U76, INTS, 6, Rue
A.Cabanef, 75015 Paris, FRANCE
FEATURES
location/Qualifiers
1..2790
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p35"
/clone="RhXIII"
/issue_type="bone marrow"
/issue_1ib="human bone marrow cDNA"
17..1270
/codon_start=1
/product="Rhesus polypeptide XIII"
/protein_id="CAA4811.1"
/db_xref="GI:36047"
/db_xref="GDB:002161"

```

CDS

```

/translation="MSSKPRSVRCPLMALTEALILFYEFTHYDASLEDOGL
VASVYGGDLTYMAIGLPLTSFRSRMSVAFPMALGVOMAILDGLFSOP
SGKIVITLFLRLATMSALSVLISVDVAGVLAOLVVMVLVEVTLRMVTSNT
FTVDHMMMHVYFAAYFGLSVAMCPEKPIPECTBEDDQVATTPSIALMALPLMI
FWPSPSALLSRPIERKAVPNTYAVAVSVTAISGSLHHPGKISKYVHSAVLA
GVAVGTCHLIPSPMLAMVGLVAGLISVGAAYLRPCCNRVIGIPIHSSIMGVNFSL
IGLSEILYIVLVLDVAGNMIIGPVVLISIGLSIAIVIALTSLTGLTLNLKI
WAPHEAKYFDQVFWKPHLAVGF"

```

ORIGIN

```

Query Match 99.9%; Score 1252.4; DB 9; Length 2790;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGAGCTTAAGTACCCGCGGTCTGTCCGCGCTGCTGCCCTCTGAGCCCTTAACACTG 60
Db 17 ATGAGCTTAAGTACCCGCGGTCTGTCCGCGCTGCTGCCCTCTGAGCCCTTAACACTG 76
Qy 61 GAAGAGCTCATTTCTCTTCTATTTTTCCTTATGAGAGCTTCTTGAAGAT 120
Db 77 GAAGAGCTCATTTCTCTTCTATTTTTCCTTATGAGAGCTTCTTGAAGAT 136
Qy 121 CAAAAGGGCTGCGGATCCCTATCAAGTTGGCCAGATGAGCCGATGAGGCGCAT 180
Db 137 CAAAAGGGCTGCGGATCCCTATCAAGTTGGCCAGATGAGCCGATGAGGCGCAT 196
Qy 181 GCGTTGGCTTCTGACCTGAGATTTCCGAGACACAGCTGAGAGCTGTGGCTTCAAC 240
Db 197 GCGTTGGCTTCTGACCTGAGATTTCCGAGACACAGCTGAGAGCTGTGGCTTCAAC 256
Qy 241 CTCTTCATGCTGGGCTGTGGAGTGGAGTGGCAATCTGCTGAGACGCTTCTGAGCCAG 300
Db 257 CTCTTCATGCTGGGCTGTGGAGTGGAGTGGCAATCTGCTGAGACGCTTCTGAGCCAG 316
Qy 301 TTCCCTTCTGGAAGGTGATCATCACTGTTCAGTATTTGGCTGCGACCATGAGTGT 360
Db 317 TTCCCTTCTGGAAGGTGATCATCACTGTTCAGTATTTGGCTGCGACCATGAGTGT 376
Qy 361 TTGTCGCTGCTGATCTCAGTGAATGCTGTTTGGGGAAGTCAACTTGGCGAGTTGGT 420
Db 377 TTGTCGCTGCTGATCTCAGTGAATGCTGTTTGGGGAAGTCAACTTGGCGAGTTGGT 436
Qy 421 GTGATGCTGTGTGAGGAGTGAAGCTTTAGGCAACTGAGATGAGTCAATGATATTC 480
Db 437 GTGATGCTGTGTGAGGAGTGAAGCTTTAGGCAACTGAGATGAGTCAATGATATTC 496

```



```

Qy 601 ACAGCAAGATACCAAGTTTGTCTGCCATGCTGGGGCCCTCTTCTTGATGATGTTCTGG 660
    |||
Db 601 ACAGCAAGATACCAAGTTTGTCTGCCATGCTGGGGCCCTCTTCTTGATGATGTTCTGG 660
Qy 661 CCAAGTTTCAACTCTGCTGCTGAGAGTCCATGAAAGAGAAATGCCGTGTTCAAC 720
    |||
Db 661 CCAAGTTTCAACTCTGCTGCTGAGAGTCCATGAAAGAGAAATGCCGTGTTCAAC 720
Qy 721 ACCTACTATGCTGTAGCAGTCAAGCTGTGTGACAGCCATCTCAAGGTCATCTTGGCTAC 780
    |||
Db 721 ACCTACTATGCTGTAGCAGTCAAGCTGTGTGACAGCCATCTCAAGGTCATCTTGGCTAC 780
Qy 781 CCCCAAGGGAAGATCAGCAAGACTTATGTGACAGTGCCTGTGGCAGAGCCGTGCT 840
    |||
Db 781 CCCCAAGGGAAGATCAGCAAGACTTATGTGACAGTGCCTGTGGCAGAGCCGTGCT 840
Qy 841 GTGGGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
    |||
Db 841 GTGGGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Qy 901 GCTGGGCTGATCTCCGTGGGGAGCCAAAGTACTGCGGGGTGTGTAAACCGAGTCTG 960
    |||
Db 901 GCTGGGCTGATCTCCGTGGGGAGCCAAAGTACTGCGGGGTGTGTAAACCGAGTCTG 960
Qy 961 GGGATTCCCAAGCTCATCAATGAGGCTCAACTTCACTTGTGGTGTGGTGTGGAGAG 1020
    |||
Db 961 GGGATTCCCAAGCTCATCAATGAGGCTCAACTTCACTTGTGGTGTGGTGTGGAGAG 1020
Qy 1021 ATCATCTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
    |||
Db 1021 ATCATCTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Qy 1081 TTCAGGCTCTCTCAGAGATGGGGAATTCAGCTTGGCCATGCTGATGATCTTCAAGTCT 1140
    |||
Db 1081 TTCAGGCTCTCTCAGAGATGGGGAATTCAGCTTGGCCATGCTGATGATCTTCAAGTCT 1140
Qy 1141 GGTCTCTGAGAGGTTTCTCTTAATCTTAAATATGAAAGAAAGCACTGATGAGGCTTAA 1200
    |||
Db 1141 GGTCTCTGAGAGGTTTCTCTTAATCTTAAATATGAAAGAAAGCACTGATGAGGCTTAA 1200
Qy 1201 TATTTTATGACCAAGTTTCTGGAAGTTCTCTGATTTGGCTGTGATTTTAA 1254
    |||
Db 1201 TATTTTATGACCAAGTTTCTGGAAGTTCTCTGATTTGGCTGTGATTTTAA 1254

RESULT 8
AB018967 1254 bp mRNA linear PRI 03-JUL-1999
LOCUS AB018967
DEFINITION Homo sapiens RhDva(TO) mRNA for Rh blood group D antigen (RHD),
complete cds.
ACCESSION AB018967
VERSION AB018967.1 GI:5360244
KEYWORDS RhDva(TO); Rh blood group D antigen (RHD).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE Hyodo, H., Ishikawa, Y., Kashiwase, K., Ogawa, A., Watanabe, Y.,
AUTHORS Tanueyama, H., Toyoda, C., Uchikawa, M., Akaza, T. and Fujii, T.
TITLE Polymorphisms of RhDva in Japanese
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1254)
AUTHORS Uchikawa, M., Hyodo, H. and Ishikawa, Y.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1998) Hi-Jonobu Hyodo, Japanese Red Cross Central
Blood Center, Research, 4-1-31, Hi-roo, Shibuya, Tokyo 150-0012,
Japan (E-mail:hyodo@ha.cbc.jrc.or.jp, Tel:+81-3-5485-6009,
Fax:+81-3-3406-7892)
FEATURES
            location/Qualifiers
            1..1254
            /organism="Homo sapiens"
            /mol_type="mRNA"

```

```

/isolate="TO"
/db_xref="taxon:9606"
1..1254
/gene="RhDva(TO)"
CDS
1..1254
/gene="RhDva(TO)"
/note="RhD variant RhDva(TO)"
/codon_start=1
/product="Rh blood group D antigen (RHD)"
/protein_id="BAA81900.1"
/db_xref="GI:5360245"
/translation="MSSKYPKPRVRLPLMALTLBAALLPYFTTHDASLEDKGL
VASVOGDLPTWMAAIGLSPFLSSRRHSVAFNLEMLALGVWAILDGLSOP
SGKRVITLPSIRLATMSALSVLISVDVAVGKNLAQVVMVLEVTALGNLMTVSNL
PNDYHNMNMHIYVFAAYPLGSLVAMCLPELPEGEDDKQDQATIPSLAMGALFLM
PWFNSALRSDPIQRKNAMFNTYVAIVAVSVTAISGSLAHPQKISKTYHSAVLA
GGVAVGTSCHLIPSPMLAMVLGLVAGLISVGAKYLLPCGNRNVGIGIPHSISNGVRS
LGLLGEIIVLVLLVDVAGNMTIGFVULSISBLSLAIVLAALTSGLLTGLMLTKI
MKAPHEAKYFDQVFWKPHLAVGF"

```

```

ORIGIN
Query Match 99.7%; Score 1250.8; DB 9; Length 1254;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 ATGAGCTCTAAGTACCCGGGCTGTGTGCGGCGCTGCTGCGCCCTTGGGCCCTTAACACTG 60
    |||
Db 1 ATGAGCTCTAAGTACCCGGGCTGTGTGCGGCGCTGCTGCGCCCTTGGGCCCTTAACACTG 60
Qy 61 GAAGCAGCTCTGATCTCTCTTCTATTTTAAATCCCACTAAGAGCTTCTTGAAGAGAT 120
    |||
Db 61 GAAGCAGCTCTGATCTCTCTCTTCTATTTTAAATCCCACTAAGAGCTTCTTGAAGAGAT 120
Qy 121 CAAAAGGGCTGTGGCATCTCATTAAGTTGGCCAAAGTCTGACCGTGATGGGCCCAT 180
    |||
Db 121 CAAAAGGGCTGTGGCATCTCATTAAGTTGGCCAAAGTCTGACCGTGATGGGCCCAT 180
Qy 181 GGGCTGGGCTTCCATCTCGACGTTTCCGAGACACAGCTGGAGCACTGGGCTTTCAAC 240
    |||
Db 181 GGGCTGGGCTTCCATCTCGACGTTTCCGAGACACAGCTGGAGCACTGGGCTTTCAAC 240
Qy 241 GGTCTGGGCTTCTCACTCAAGTTTCCGAGACACAGCTGGAGCACTGGGCTTTCAAC 240
    |||
Db 241 GGTCTGGGCTTCTCACTCAAGTTTCCGAGACACAGCTGGAGCACTGGGCTTTCAAC 240
Qy 241 CTCTTCATGCTGGCGCTTGTGTGAGTGGCAATCTCTGCTGACCGCTTCTGAGCAG 300
    |||
Db 241 CTCTTCATGCTGGCGCTTGTGTGAGTGGCAATCTCTGCTGACCGCTTCTGAGCAG 300
Qy 301 TTCCCTTGGGAAAGGTGATCATCACTGTTCATGATTTGGGCTGGCCACCATGATGCT 360
    |||
Db 301 TTCCCTTGGGAAAGGTGATCATCACTGTTCATGATTTGGGCTGGCCACCATGATGCT 360
Qy 361 TTGTGGTGTGATCTGATGATGATGCTGTCTTGGGGAAGTCAACTTGGGCGAGTTGGTG 420
    |||
Db 361 TTGTGGTGTGATCTGATGATGATGCTGTCTTGGGGAAGTCAACTTGGGCGAGTTGGTG 420
Qy 421 GTGATGGTGTGTGTGGAGTGAACAGCTTTAGGCAACCTTGAGATGGTCATCAATAATATC 480
    |||
Db 421 GTGATGGTGTGTGTGGAGTGAACAGCTTTAGGCAACCTTGAGATGGTCATCAATAATATC 480
Qy 481 TTCAACAACAAGTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
    |||
Db 481 TTCAACAACAAGTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Qy 541 CTGTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 600
    |||
Db 541 CTGTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 600
Qy 601 ACAGCAAGATACCAAGTTTGTCTGCCATGCTGGGGCCCTCTTCTTGATGATGTTCTGG 660
    |||
Db 601 ACAGCAAGATACCAAGTTTGTCTGCCATGCTGGGGCCCTCTTCTTGATGATGTTCTGG 660
Qy 661 CCAAGTTTCAACTCTGCTGCTGAGAGTCCATGAAAGAGAAATGCCGTGTTCAAC 720
    |||
Db 661 CCAAGTTTCAACTCTGCTGCTGAGAGTCCATGAAAGAGAAATGCCGTGTTCAAC 720

```

[illegible]

/protein_id="AA075125.1"
/db_xref="GI:25573108"
/translation="MSKSGPRSVRCLPLMALTLLEAAILLPEFTTHYDASEDQGL
VASYVGODLTWMAALGIGLITSPRRHSSSAFNLPMALGQVAMAILDGLSOP
SGGVVTLTPELRATMSALSVLSVDAYVAKVLAOLVNVVALVEVATLRVAVINSI
PNTDYPANMMHAIYVFAAYRGLSVAKCLPRLDPEGTEDKOOTATPISLNAIIGALPLMI
FMSVNSALLRSPDIEKKNAFNTIYVAANSVTATIGSSLSLAHQGKSKITYHSAYLA
GGVAVSGSCHLIPSPMLAWYGVAGISVGAKNIPGCCNRVLDGTPHSSIMGNFSL
LGLALGELIYLVLLVLTGAGNMGITGVLLSIGELSLAIVALTSELGLTGLTLNFI
WKAPEHKAFFDDQVFWKPHLAVGP"

Query Match	99.7%	Score 1250.8;	DB 9;	Length 1480;
Best Local Similarity	99.8%	Pred. No. 0;		
Matches 1252; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

QY	1	ATGAGGCTCTAAGAACCCGCGGCTGTCTCCGCGCTGTCCCTGTGGGCTTAACACTG	60
Db	44	ATGAGGCTCTAAGAACCCGCGGCTGTCTCCGCGCTGTCCCTGTGGGCTTAACACTG	103
QY	61	GAAGCAGCTCATTTCTCTCTTCTATTTTTTTTAAACCATATGACGTTCTTAAAGAGT	120
Db	104	GAAGCAGCTCTCATTTCTCTCTTCTATTTTTTTTAAACCATATGACGTTCTTAAAGAGT	163
QY	121	CAAAAGGGGCTCGGCGCATCTTATCAAGTTGGCCAAAGATCGAACCGATGGCGGCATT	180
Db	164	CAAAAGGGGCTCGGCGCATCTTATCAAGTTGGCCAAAGATCGAACCGATGGCGGCATT	223
QY	181	GGCTTGGGCTTCTCTACCTCGAGTTTCCGAGACACACGCTGGAGCAGTGTGGCTTCAAC	240
Db	224	GGCTTGGGCTTCTCTACCTCGAGTTTCCGAGACACACGCTGGAGCAGTGTGGCTTCAAC	283
QY	241	CTCTTCAATGCTGGCGCTTGGTGTGACGTGGGCAATCTGCTGACCGGCTTCTGAGCGAG	300
Db	284	CTCTTCAATGCTGGCGCTTGGTGTGACGTGGGCAATCTGCTGACCGGCTTCTGAGCGAG	343
QY	301	TTCCCTTCTGGGAAGGTGTCATCACACTTTTCAGTATTCGCGTGGCCACCATGATGCT	360
Db	344	TTCCCTTCTGGGAAGGTGTCATCACACTGTTCAGTATTCGCGTGGCCACCATGATGCT	403
QY	361	TTGTCCGTGCTGATCTTCAGTGGATGCTGTCTTGGGGAAAGTCAACTTGGCGCAGTTGGTG	420
Db	404	TTGTCCGTGCTGATCTTCAGTGGATGCTGTCTTGGGGAAAGTCAACTTGGCGCAGTTGGTG	463
QY	421	GTTATGCTGCTGGTGGAGGTGACAGACTTTAAGGCAACTGAGAGATGTCATCAGTAATATC	480
Db	464	GTTATGCTGCTGGTGGAGGTGACAGACTTTAAGGCAACTGAGAGATGTCATCAGTAATATC	523
QY	481	TTCAACACAGACTAACCATATGATACATCTACGCTGTTCGACGCTATATTTGGG	540
Db	524	TTCAACACAGACTAACCATATGATACATCTACGCTGTTCGACGCTATATTTGGG	583
QY	541	CTGTCTGTGGCTTGTGTGCTCCCAAGCCTTCAACCCGAGGGAAACGAGATTAAGATCAG	600
Db	584	CTGTCTGTGGCTTGTGTGCTCCCAAGCCTTCAACCCGAGGGAAACGAGATTAAGATCAG	643
QY	601	ACAAGACAAGATACCAGATTTGTCTMGCCATGCTGGGCGCCCTCTTCTTGATGATTTCTGG	660
Db	644	ACAAGACAAGATACCAGATTTGTCTMGCCATGCTGGGCGCCCTCTTCTTGATGATTTCTGG	703
QY	661	CCAAGTTTCAACTCTGTCTCTGTGAGAAAGTCCAATCGAAAGAAAGAAATGCGCTGTCAAC	720
Db	704	CCAAGTTTCAACTCTGTCTCTGTGAGAAAGTCCAATCGAAAGAAAGAAATGCGCTGTCAAC	763
QY	721	ACCTACTATGCTGTAGAGATCAGGCTGTGTACAGCCATCTCAAGGTATATCTTGGCTTAC	780
Db	764	ACCTACTATGCTGTAGAGATCAGGCTGTGTACAGCCATCTCAAGGTATATCTTGGCTTAC	823
QY	781	CCCCAAGGGAAGATCAGCAAGACTTATATGTCAACAGTGCAGTGTTCGACGAGCGCTGCT	840
Db	824	CCCCAAGGGAAGATCAGCAAGACTTATATGTCAACAGTGCAGTGTTCGACGAGCGCTGCT	883
QY	841	GTGGGTACCTGTGTACCTGATCTCTTCTTCGCTGCTTGCACATGATGTGTGTGTG	900

Db 884 GTGGGTAACCTGGTACCTGATCCCTTCCCTGCTGGCCATGAGTGGCTGTTGG 943
QY 901 GCTGGGCTGATCTCCGTCGGGGAGCCAGATACCTCCGGGGCTTGTAAACGAGTGG 960
Db 944 GCTGGGCTGATCTCCGTCGGGGAGCCAGATACCTCCGGGGCTTGTAAACGAGTGG 1003
QY 961 GGGATTTCCCAACGCTTCATCATGAGGCTTCAACCTTCAAGCTTGTGGGCTTGGAGAG 1020
Db 1004 GGGATTTCCCAACGCTTCATCATGAGGCTTCAACCTTCAAGCTTGTGGGCTTGGAGAG 1063
QY 1021 ATGATCTACATTTGTGGGCTTGTGGGCTTGTAAACCTTGGAGCCGGCAATGGCATTTGGC 1080
Db 1064 ATGATCTACATTTGTGGGCTTGTGGGCTTGTAAACCTTGGAGCCGGCAATGGCATTTGGC 1123
QY 1081 TTCCAGGCTCTCTCTGAGCATTTGGGGAACTGAGCTTGGCCATGATAGCTTCACGCT 1140
Db 1124 TTCCAGGCTCTCTCTGAGCATTTGGGGAACTGAGCTTGGCCATGATAGCTTCACGCT 1183
QY 1141 GGCTCTGACAGCTTGTGCTCTTAAATCTTAAATATGGAAGCACTCATGAGGCTTAA 1200
Db 1184 GGCTCTGACAGCTTGTGCTCTTAAATCTTAAATATGGAAGCACTCATGAGGCTTAA 1243
QY 1201 TATTTGATGACCAAGTTTCTGGAAGTTCTCTCTCTTGGCTTGGATTTTAA 1254
Db 1244 TATTTGATGACCAAGTTTCTGGAAGTTTCTCTCTCTTGGCTTGGATTTTAA 1297

RESULT 10
AF510070 1480 bp mRNA linear PRI 27-NOV-2002
LOCUS Homo sapiens Rhesus blood group D antigen D(674) variant mRNA,
DEFINITION complete cds.
ACCESSION AF510070
VERSION AF510070.1 GI:25573111
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1480)
AUTHORS Noizat-Pirenne, F., Lee, K., Le Pennec, P. Y., Simon, P., Kazup, P.,
Bachir, D., Rouzaud, A. M., Rousset, M., Juszczak, G., Menanteau, C.,
Rouger, P., Koch, R., Cartlon, J. P. and Ansart-Pirenne, H.
Rare RHCE phenotypes in black individuals of Afro-Caribbean origin:
identification and transfusion safety
Blood 100 (12), 4223-4231 (2002)
JOURNAL PUBMED 12393640
REFERENCE 2 (bases 1 to 1480)
AUTHORS Noizat-Pirenne, F., Cartlon, J. P. and Ansart-Pirenne, H.
TITLE Direct Submission
JOURNAL Submitted (08-MAY-2002) CNRGS, INTS, 20 rue Bouvier, Paris 75011,
France
FEATURES
source
Location/Qualifiers
1. 1480
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/note="Isolated from Black individual of Afro-Caribbean
origin"
44. 1297
/note="RHD"
/product="Rhesus blood group D antigen D(674) variant"
/protein_id="AA075112.1"
/db_xref="GI:25573112"
/translation="MSKYPRSVRCLPLMALTLLENALTLFFPYTHYDASLEDKGL
VASVQGDLTVAALIGFLTSFRRHSVSAFNLFWALGVOMATILIGFSGFP
SGKVVITLSIRLATWSALSVLSVDAVGVKNLAOLVVMALGVOMATILIGFSGFP
FNTDYHMMNHIVFAVYFGLSVAMCLPKPLPEGTEDDQDTTSTISAMGALFLMI
FMPSEFALLRSPIERKNVENVYVAVNSVTATSSGSLAHPOKSTIKTVHSAVLA
GGVAVGTSCHLIPSPILAMVGLVAGLISVGAKTLPCCNKVLSIPSSITGVNFSL
LGLLEIITIVLIVDTVGAGNGMTGFOVLISGLSLAIVALTSLTLGLTLNLKI

CDS

ORIGIN WKAPHEAKTYDDOVFWKPHLAVGP"
Query Match 99.7%; Score 1250.8; DB 9; Length 1480;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGCTCTAAGTACCGGCGGTCTGTCGGGCTGCTCCCTCTGGAGCCCTAACACTG 60
Db 44 ATGAGCTCTAAGTACCGGCGGTCTGTCGGGCTGCTCCCTCTGGAGCCCTAACACTG 103
QY 61 GAAGAGCTCTCATCTCCCTCTTATTTTAAACCATAGACGCTTCTTAAGAT 120
Db 104 GAAGAGCTCTCATCTCCCTCTTATTTTAAACCATAGACGCTTCTTAAGAT 163
QY 121 CAAAGGGGCTGTGAGCATCTATCAAGTTGGCAAGATCTGACCGTGAATGAGCAT 180
Db 164 CAAAGGGGCTGTGAGCATCTATCAAGTTGGCAAGATCTGACCGTGAATGAGCAT 223
QY 181 GGCTTGGGCTTCTCTCACTTCCGAGTTTCCGAGACACAGCTGAGCATGTGCTTCAAC 240
Db 224 GGCTTGGGCTTCTCTCACTTCCGAGTTTCCGAGACACAGCTGAGCATGTGCTTCAAC 283
QY 241 CTCTTATGCTGAGCGCTTGTGTGAGTGGGCAATCTGCTGAGGCTTCTGAGCCAG 300
Db 284 CTCTTATGCTGAGCGCTTGTGTGAGTGGGCAATCTGCTGAGGCTTCTGAGCCAG 343
QY 301 TTCCCTTCTGGGAGAGTGTCTATCAACTGTTCACTATTCGGCTGGCCACATGAGTGT 360
Db 344 TTCCCTTCTGGGAGAGTGTCTATCAACTGTTCACTATTCGGCTGGCCACATGAGTGT 403
QY 361 TTGTGCTGCTGATCTCACTGATGCTGCTTGGGAGGTCACTTGGCCAGTGTGTG 420
Db 404 TTGTGCTGCTGATCTCACTGATGCTGCTTGGGAGGTCACTTGGCCAGTGTGTG 463
QY 421 GTGATGTGCTGTGGAGGTGACAGCTTTAGGCAACTGAGATGTCATGATATATC 480
Db 464 GTGATGTGCTGTGGAGGTGACAGCTTTAGGCAACTGAGATGTCATGATATATC 523
QY 481 TTCAACACAGACTTACCAATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 524 TTCAACACAGACTTACCAATGATGATGATGATGATGATGATGATGATGATGATG 583
QY 541 CTGTCTGTGCTGTGTGCTGCTCAAGGCTTCAACCGAGGGAAGAGATTAAGATCAG 600
Db 584 CTGTCTGTGCTGTGTGCTGCTCAAGGCTTCAACCGAGGGAAGAGATTAAGATCAG 643
QY 601 ACAGCAAGATACCAAGTTGTGTGCAATGCTGGGCGCTTCTTGTGATGTCTGG 660
Db 644 ACAGCAAGATACCAAGTTGTGTGCAATGCTGGGCGCTTCTTGTGATGTCTGG 703
QY 661 CCAAGTTCACTGCTGCTGTGAGAGTCCATGGAAGTCCATGGAAGAAATGCGCGTTCAC 720
Db 704 CCAAGTTCACTGCTGCTGTGAGAGTCCATGGAAGTCCATGGAAGAAATGCGCGTTCAC 763
QY 721 ACTTACTATGTGTAGACATGACGCTGTGAGACAGCATCTCAGGGTCACTTGGCTCAC 780
Db 764 ACTTACTATGTGTAGACATGACGCTGTGAGACAGCATCTCAGGGTCACTTGGCTCAC 823
QY 781 CCCCAGGAAAGATCACCAAGATCTTATGTGCAAGTGTGCTGTTGGAGAGGCGTGGCT 840
Db 824 CCCCAGGAAAGATCACCAAGATCTTATGTGCAAGTGTGCTGTTGGAGAGGCGTGGCT 883
QY 841 GTGGGTAACCTGAGTCACTGATCCCTTCCGAGCTTGCATGAGTGTGCTGTTGG 900
Db 884 GTGGGTAACCTGAGTCACTGATCCCTTCCGAGCTTGCATGAGTGTGCTGTTGG 943
QY 901 GCTGGGCTGATCTTCCTGCGGGAGGCAAGTACTTCCGGGGTGTGTAACGAGTGGT 960
Db 944 GCTGGGCTGATCTTCCTGCGGGAGGCAAGTACTTCCGGGGTGTGTAACGAGTGGT 1003
QY 961 GGGATTTCCCAACGCTTCATCATGAGGCTTCAACCTTCAAGCTTGTGGGCTTGGAGAG 1020

Db		1004	GGGATTTCCCCACAGCTCATCATGAGGCATCAACAATTCAAGCTTGCTGGGTCTGCTTGGAAAG	1063
Qy		1021	ATCATCTTAACATTTGTGCTGCTGTGCTTGAATACCGTCGGAGCCGGCAATGAGCATGATTTGGC	1080
Db		1064	ATCATCTTAACATTTGTGCTGCTGTGCTTGAATACCGTCGGAGCCGGCAATGAGCATGATTTGGC	1123
Qy		1081	TTCCAGGCCCTCCATCAGATTTGGGGAACTCAGCTTGGGCATTCGTGATGCTCAACGCTCT	1140
Db		1124	TTCCAGGCCCTCCATCAGATTTGGGGAACTCAGCTTGGGCATTCGTGATGCTCAACGCTCT	1183
Qy		1141	GCTCTCTGACAGGTTTTGCTCTCTAAATCTTAAATAATGAAAAGCACCTCATGAGGCTAA	1200
Db		1184	GCTCTCTGACAGGTTTTGCTCTCTAAATCTTAAATAATGAAAAGCACCTCATGAGGCTAA	1243
Qy		1201	TATTTTGAATGACCAGATTTTCTGGAAGTTTCTCATTTTGCGCTGTTGATTTTAA	1254
Db		1244	TATTTTGAATGACCAGATTTTCTGGAAGTTTCTCATTTTGCGCTGTTGATTTTAA	1297
RESULT	11			
LOCUS	AJ784309			
DEFINITION	Homo sapiens partial RHD gene for rhesus blood group D antigen,			
ACCESSION	AJ784309	1251 bp	DNA	linear
VERSION	AJ784309.1			PRI 25-JUL-2004
KEYWORDS	RHD gene; rhesus blood group D antigen; weak D allele.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
REFERENCE	Doeschner A., Wagner F., Schunter F. and Petershofen E.K. Four new RHD alleles with previously unknown polymorphism Unpublished 2 (bases 1 to 1251) Doeschner A. Direct Submission Submitted (13-JUL-2004) Doeschner A., Molecular Diagnostics, German Red Cross Blood Transfusion Center, Brandenburger Str. 21, D-26133 Oldenburg, GERMANY Related Sequence: L08429.			
JOURNAL				
FEATURES				
source				
gene				
CDS				
exon				
variation				

ORIGIN	/gene="RHD" /note="single nucleotide polymorphism, SNP" /replace="C"	Query Match Best Local Similarity Matches 1249, Conservative	99.5%; 99.8%; 0;	Score 1247.8; Pred. No. 0; 0;	DB 9;	Length 1251; 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGCTGCTGCCCTCTTGCGCCCTTAACACTG	60							
DB	1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGCTGCTGCCCTCTTGCGCCCTTAACACTG	60							
QY	61 GAACAGCTCATATTCCTCTTCTATTTTTTTTACCACTATGACGCTTCTTGAGGAT	120							
DB	61 GAACAGCTCATATTCCTCTTCTATTTTTTTTACCACTATGACGCTTCTTGAGGAT	120							
QY	121 CAAAGGGGCGCTGAGCATCTTATCAAGTTGGCCAAAGATCTGACCGTATGGCGCCATT	180							
DB	121 CAAAGGGGCGCTGAGCATCTTATCAAGTTGGCCAAAGATCTGACCGTATGGCGCCATT	180							
QY	181 GGGTTGGGCTTCTCACTCGAGTTTCCGAGACAACAGCTGAGACAATGTGGCTTCAAC	240							
DB	181 GGGTTGGGCTTCTCACTCGAGTTTCCGAGAGACAACAGCTGAGACAATGTGGCTTCAAC	240							
QY	241 CTCTTCAATGCTGGGCTTGGTGTGACAGGGGCATTCGTGCTGACGCGTCTTGAGCCAG	300							
DB	241 CTCTTCAATGCTGGGCTTGGTGTGACAGGGGCATTCGTGCTGACGCGTCTTGAGCCAG	300							
QY	301 TTCCCTTCTGGGGAAGGTGTCATCACTGTTCAGTATTCGGCTGGCCACATGAGTCT	360							
DB	301 TTCCCTTCTGGGGAAGGTGTCATCACTGTTCAGTATTCGGCTGGCCACATGAGTCT	360							
QY	361 TTGTGCGTCTGATCTCAGTGAATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGGTG	420							
DB	361 TTGTGCGTCTGATCTCAGTGAATGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGGTG	420							
QY	421 GTATGTGCTGCTGGTGGAGGTGACAGCTTTAGGCAACCTGAGAGTGGTATCAGTAATATC	480							
DB	421 GTATGTGCTGCTGGTGGAGGTGACAGCTTTAGGCAACCTGAGAGTGGTATCAGTAATATC	480							
QY	481 TTCAACACAGACTCAACATGATGATGACATCTACGTTCGACGCTATTTTGGG	540							
DB	481 TTCAACACAGACTCAACATGATGATGACATCTACGTTCGACGCTATTTTGGG	540							
QY	541 CTGTCTGTGGCTGTGGTCTTCCAAAGCTCTTACCCGAGGGAAACGAGATTAAGATCAG	600							
DB	541 CTGTCTGTGGCTGTGGTCTTCCAAAGCTCTTACCCGAGGGAAACGAGATTAAGATCAG	600							
QY	601 ACACCAACGATACCAAGTTGTGTGCAATGCGGGCGCCCTTGTGTGATGTTCTGG	660							
DB	601 ACACCAACGATACCAAGTTGTGTGCAATGCGGGCGCCCTTGTGTGATGTTCTGG	660							
QY	661 CCAAGTTTCAACTCTGCTCTGCTGAGAGATCCAAATCGAAAGGAAGATGCGTGTTCAC	720							
DB	661 CCAAGTTTCAACTCTGCTCTGCTGAGAGATCCAAATCGAAAGGAAGATGCGTGTTCAC	720							
QY	721 ACCTACTATGTCTGTAGCAGTCAAGGTGTGACAGCATCTCAAGGATCATCTTGGCTCAC	780							
DB	721 ACCTACTATGTCTGTAGCAGTCAAGGTGTGACAGCATCTCAAGGATCATCTTGGCTCAC	780							
QY	781 CCCCAGGGAAGATCAGCAAGACTTAATGTGACAGTGGGAGTTTGGCAGAGGCGTGGCT	840							
DB	781 CCCCAGGGAAGATCAGCAAGACTTAATGTGACAGTGGGAGTTTGGCAGAGGCGTGGCT	840							
QY	841 GTGGGTACTGTGTCACTGATCCCTTCTCCGTGGCTTGGCATGTGCTGGCTTGTG	900							
DB	841 GTGGGTACTGTGTCACTGATCCCTTCTCCGTGGCTTGGCATGTGCTGGCTTGTG	900							
QY	901 GCTGGGCTGATCTCCGTGGGGGAGCCAAATGACCTGCGGGGATTTTGAACGAATGCTG	960							
DB	901 GCTGGGCTGATCTCCGTGGGGGAGCCAAATGACCTGCGGGGATTTTGAACGAATGCTG	960							

Db	901	GCTGGGCGTGAATCTCCGTGGGGGGAGCAAGTACCTGCCGGGGGTGTGTAAACCAAGTGTG	960
Qy	961	GGGATTTCCCAACAGTCTCCATCATAGGGGTACAACTTCAGCTTGCTGGGTCTGCTTGGAGAG	1020
Db	961	GGGATTTCCCAACAGTCTCCATCATAGGGGTACAACTTCAGCTTGCTGGGGTCTCTTGGAGAG	1020
Qy	1021	ATCATCTACATTTGTGCTCTGGTGGTTATATACCGTCGGAGCCGGCAATGGCATGTATGGC	1080
Db	1021	ATCATCTACATTTGTGTGTGTGCTGGTGTATACCGTCGGAGCCGGCAATGGCATGTATGGC	1080
Qy	1081	TTCCAGGTCTCTCTCAGCATTTGGGGAACTCAGCTTGGCCATCTGTGATAGCTCTCACTCT	1140
Db	1081	TTCCAGGTCTCTCTCAGCATTTGGGGAACTCAGCTTGGCCATCTGTGATAGCTCTCACTCT	1140
Qy	1141	GGTCTCTCGAAGGTTTGTCTCTTAAATCTTAAATATGAAAGCACCTCATAGAGGCTAAA	1200
Db	1141	GGTCTCTCGAAGGTTTGTCTCTTAAATCTTAAATATGAAAGCACCTCATAGAGGCTAAA	1200
Qy	1201	TATTTTATGACCAAGTTTCTTGGAACTTCTCTATTTGGCTGTGTGATTT	1251
Db	1201	TATTTTATGACCAAGTTTCTTGGAACTTCTCTATTTGGCTGTGTGATTT	1251

	RESULT 13
AJ784311	
LOCUS	AJ784311
DEFINITION	Homo sapiens partial RHD gene for rhesus blood group D antigen,
VERSION	wkD allele, exon 4.
KEYWORDS	
SOURCE	AJ784311.1 GI:50657036
ORGANISM	RHD gene; rhesus blood group D antigen. Homo sapiens (human) Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 Doescher A., Wagner F., Schunter F. and Peterhofen, E.K. Four new RHD alleles with previously unknown polymorphism Unpublished 2 (bases 1 to 1251)
AUTHORS	Doescher, A.
TITLE	Direct Submision
JOURNAL	Submitted (19-JUL-2004) Doescher A.. Molecular Diagnostics, German Red Cross Blood Transfusion Centre, Brandenburger Str. 21, D-26133 Oldenburg, GERMANY
COMMENT	Related sequence: L08429.
FEATURES	Location/Qualifiers
source	1..1251
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/isolate="19/2004"
	/db_xref="taxon:9606"
	/chromosome="1"
	/map="tp36.2-p34"
	/cell_type="erythrocyte"
	/tissue_type="peripheral blood"
gene	1..1251
	/gene="RHD"
CDS	<1..>1251
	/gene="rhd"
	/codon_start=1
	/product="rhesus blood group D antigen"
	/allele="weak_D"
	/protein_id="CAH04921.1"
	/db_xref="GI:50657037"
	/translation="MSKYPERSVRCLPLMALTLLEAALLILFFFTHYDSLDEQGLVASYOVGDLLTMAAIGCGELTSFRRHSWSVAFLMLAGVAMAILLDGFSOPPEASGVITLFSRLATMSALSVLIADVAVGNLAOLVVMTLVETVLGNLRWITSNIRFNTRYNNMMHTIYMPAAIEGYISVMCLPKLPDETEKDORATI PSLSNMGALFLMMFWSPNSALLSPISPIERKANVTYTAANAVSVTYAISSGSLLHPGAKISKTYSHAYILA GGVAVGTSCHLIPSPMLAWNLGTVALGISVGARYLPPCCNRVIGISHSSIMGFNFSL KGLLGEEITYIVLVDLTVGAGNGMIIGPVLVISIELSLAIVIALITSGLLTGELTLNKIKWKAITHAEYFPDDQVFWEKPRLAVGF"

exon	1. .1251	/gene="RHD"
variation	520	/number=4
	/gene="RHD"	
	/note="single nucleotide polymorphism, SNP"	
	/replace="g"	
ORIGIN		
Query Match	99.5%;	Score 1247.8 ; DB 9 ; Length 1251 ;
Best Local Similarity	99.8%;	Pred. No. 0 ;
Matches 1249 ;	Conservative 0 ;	Mismatches 2 ; Indels 0 ; Gaps 0 ;
QY	1	ATGAGCTCTAAGTACCCGCGGTCTGTCCGGGCTGCTGCCCTCTGGGCGCTTAACCTG 60
DB	1	ATGAGCTCTAAGTACCCGCGGTCTGTCCGGGCTGCTGCCCTCTGGGCGCTTAACCTG 60
QY	61	GAACAGCTCTCATCTCTCTCTTCTAATTTTAAATCCCACTATGACGCTCTCTAGAGAGT 120
DB	61	GAACAGCTCTCATCTCTCTCTTCTAATTTTAAATCCCACTATGACGCTCTCTAGAGAGT 120
QY	121	CAAAAGGGGCTGTGGCATCTCTATCAAGTTGGCCAAAGTCTGACCGTGTATGCGGCCATT 180
DB	121	CAAAAGGGGCTGTGGCATCTCTATCAAGTTGGCCAAAGTCTGACCGTGTATGCGGCCATT 180
QY	181	GGCTTGGGCTTCCTACCTCGAGTTTCCGAGACACACAGCTGGACAGTGTGGCTTCAAC 240
DB	181	GGCTTGGGCTTCCTACCTCGAGTTTCCGAGACACACAGCTGGACAGTGTGGCTTCAAC 240
QY	241	CTCTTCAATGCTGGGCTGTGTGTCAGTGGGCAATCTGCTGACAGGCTTCTGAGCCAG 300
DB	241	CTCTTCAATGCTGGGCTGTGTGTCAGTGGGCAATCTGCTGACAGGCTTCTGAGCCAG 300
QY	301	TTCCCTCTGGGAAAGTGTATCACTGTTCACTATTCGGCTGGCCACATGATGCT 360
DB	301	TTCCCTCTGGGAAAGTGTATCACTGTTCACTATTCGGCTGGCCACATGATGCT 360
QY	361	TTGTGCGGTGATCTCAGTGGAGTGTCTTGGGGAAGTCACTTGGGCGAGTGGTG 420
DB	361	TTGTGCGGTGATCTCAGTGGAGTGTCTTGGGGAAGTCACTTGGGCGAGTGGTG 420
QY	421	GTGATGTGTCTGTGTGAGTGTACAGCTTTAGGCAACCTGAGATGTCTATCAATATC 480
DB	421	GTGATGTGTCTGTGTGAGTGTACAGCTTTAGGCAACCTGAGATGTCTATCAATATC 480
QY	481	TTCAAACACAGACTACACATGACATGATGCACTTACGTTGTCAGACCTATTTTGGG 540
DB	481	TTCAAACACAGACTACACATGATGATGCACTTACGTTGTCAGACCTATTTTGGG 540
QY	541	CTGTCTGTGGCTGTGTGCTGTGCAAAAGCTCTACCGAGGGGACGGAGATTAAGATCAG 600
DB	541	CTGTCTGTGGCTGTGTGCTGTGCAAAAGCTCTACCGAGGGGACGGAGATTAAGATCAG 600
QY	601	ACAGCAACGATACCAGTTGTCTGTGCATGCTGTGGGCGCCCTCTTCTGTGATGTTCTGG 660
DB	601	ACAGCAACGATACCAGTTGTCTGTGCATGCTGTGGGCGCCCTCTTCTGTGATGTTCTGG 660
QY	661	CCAAATTCAAATCTGTCTCTGTCTGTGAAAGTCAATCGAAAGGAAGATCCGTGTTCAAC 720
DB	661	CCAAATTCAAATCTGTCTCTGTCTGTGAAAGTCAATCGAAAGGAAGATCCGTGTTCAAC 720
QY	721	ACCAACTATGTGTAGACAGTCAAGGTGTGACACCAATCTAGAGGTCACTCTTGGCTCAC 780
DB	721	ACCAACTATGTGTAGACAGTCAAGGTGTGACACCAATCTAGAGGTCACTCTTGGCTCAC 780
QY	781	CCCCAAGGGAAGTACAGCAACTTATGTGCACAGTCCGAGTGTGGCAGAGGCGTGGCT 840
DB	781	CCCCAAGGGAAGTACAGCAACTTATGTGCACAGTCCGAGTGTGGCAGAGGCGTGGCT 840
QY	841	GTGGGTACTGTGTGTACCTGATCCCTTCTCCGTGGCTTGGCAGTGTGTGGTCTTGTG 900
DB	841	GTGGGTACTGTGTGTACCTGATCCCTTCTCCGTGGCTTGGCAGTGTGTGGTCTTGTG 900

QY 901 GCTGGGCTGATCTCCGTCCGGGGAGCCAGTACCTGCGGGGGTGTGTAACCGAGTCTG 960
Db 901 GCTGGGCTGATCTCCGTCCGGGGAGCCAGTACCTGCGGGGGTGTGTAACCGAGTCTG 960
QY 961 GGGATTCGCCACAGCTCCCATCATGAGGCTACACTTCAAGCTTGGGGCTGCTTGGAGAG 1020
Db 961 GGGATTCGCCACAGCTCCCATCATGAGGCTACACTTCAAGCTTGGGGCTGCTTGGAGAG 1020
QY 1021 ATCATCTACATTGTGCTGTGTGTGATACCGTCCGAGCCGGCAATGGCAATGTTGGC 1080
Db 1021 ATCATCTACATTGTGCTGTGTGTGATACCGTCCGAGCCGGCAATGGCAATGTTGGC 1080
QY 1081 TTCAGAGTCTCTCTCCAGCATTTGGGGAACTCAGCTTGGCCATGTGATAGCTCTCAGCT 1140
Db 1081 TTCAGAGTCTCTCTCCAGCATTTGGGGAACTCAGCTTGGCCATGTGATAGCTCTCAGCT 1140
QY 1141 GGTCTCTGACAGTGTGCTCTCTAAATCTTAAATATGGAAGCACTCATGAGCTTAA 1200
Db 1141 GGTCTCTGACAGTGTGCTCTCTAAATCTTAAATATGGAAGCACTCATGAGCTTAA 1200
QY 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTT 1251
Db 1201 TATTTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTT 1251

RESULT 14

LOCUS AY449381 1251 bp mRNA linear PRI 03-DEC-2003
DEFINITION Homo sapiens rhesus blood group D antigen DFE mRNA, partial cde.
ACCESSION AY449381
VERSION AY449381.1 GI:38565510
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1251)
AUTHORS Noizat-Pirenne, F. and Ansart-Pirenne, H.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-2003) Blood Bank, EFS, 151 Avenue du Marechal de
Lattre de Tassigny, Creteil 94010, France
FEATURES
Source Location/Qualifiers
1..1251

CDS

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..>1251
/note="Weak D DFE"
/codon_start=1
/product="rhesus blood group D antigen DFE"
/protein_id="AA24078.1"
/db_xref="GI:38565511"
/translation="MSKSRPSRVRCLPLMALTLLENALILLYFTHYDASLEDQKGL
VASVQVQDLTTWMAIGLEPLISFRHSMSVAENLFMLAGVQWALIDPLSQPL
SGKVVITLPSIRLATWSALSVLISDAVAGKVALQDVVMEVETALINDLMEVSNL
FNTDYMNNMHIYFAVAGLSVAWCLPKPEDEGTCKQATLISLSMGLAFMLI
FWSPFSALIRSPILERKNVFNPTVYVAASVTAISGSLADPOKISTVYHSAVLA
GLDVAVGSCHLIPSPWLMVGLVAGLISVGAKXVLRGCNVNVLGIPIHSINGVPSL
LGLGEIIVILVILDTGAGNMTGFOVLISIGLSLAIVALTSLGLTLNLKLI
WKAPHAKYFDDQVFWKPHLAVGF"

ORIGIN

Query Match 99.5%; Score 1247.8; DB 9; Length 1251;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1249; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTCAAGTACCCGGGCTGTGTCGGCGCTGCTGCCCTCTGAGCCCTAACACTG 60
Db 1 ATGAGCTCAAGTACCCGGGCTGTGTCGGCGCTGCTGCCCTCTGAGCCCTAACACTG 60
QY 61 GAAGAGCTCTGATCTCTCTCTTATTTTATTTTACCACTATGAGGCTCTTATAGAT 120
Db 61 GAAGAGCTCTGATCTCTCTCTTATTTTATTTTACCACTATGAGGCTCTTATAGAT 120

QY 121 CAAGAAGGCTCTGAGCATCTATCAAGTTGGCAAGATCTGACCGGATGAGGCAATT 180
Db 121 CAAGAAGGCTCTGAGCATCTATCAAGTTGGCAAGATCTGACCGGATGAGGCAATT 180
QY 181 GGGTGGGCTTCTCACTGATGATTTTCCGAGACACAGCTGAGCAGTGTGCTTCAAC 240
Db 181 GGGTGGGCTTCTCACTGATGATTTTCCGAGACACAGCTGAGCAGTGTGCTTCAAC 240
QY 241 CTCTTCATGCTGAGGCTTGGTGTGCAAGTGGGCAATCTGTGGAAGGCTTCCGAGCCAG 300
Db 241 CTCTTCATGCTGAGGCTTGGTGTGCAAGTGGGCAATCTGTGGAAGGCTTCCGAGCCAG 300
QY 301 TTCCTTCTGGGAAGTGGTCAATCACTGTCAGTATTCGGCTGGCCACATGATGCT 360
Db 301 TTCCTTCTGGGAAGTGGTCAATCACTGTCAGTATTCGGCTGGCCACATGATGCT 360
QY 361 TTGTCGGTGTATCTCAGTGAATGCTGTGGGGAAGTCAATTGGCCGATTTGGTG 420
Db 361 TTGTCGGTGTATCTCAGTGAATGCTGTGGGGAAGTCAATTGGCCGATTTGGTG 420
QY 421 GTGATGTGCTGTGAGGTGACAGCTTTAGGCAACTGAGAGTGTATCTAGTAATATC 480
Db 421 GTGATGTGCTGTGAGGTGACAGCTTTAGGCAACTGAGAGTGTATCTAGTAATATC 480
QY 481 TTCAACACAGACTACCAATGAAATGATGCAATCTACGTTTGCAGCTTATTTGGG 540
Db 481 TTCAACACAGACTACCAATGAAATGATGCAATCTACGTTTGCAGCTTATTTGGG 540
QY 541 CTGTGTGTGGCTGTGTGCTGCTGCAAGGCTCTACCCGAGGGAACGAGAGATTAAGATAG 600
Db 541 CTGTGTGTGGCTGTGTGCTGCTGCAAGGCTCTACCCGAGGGAACGAGAGATTAAGATAG 600
QY 601 ACAGCAACGATACCAAGTTTGTCTCCATGCTGGGCGCTCTTCTTGTGATGTTCTGG 660
Db 601 ACAGCAACGATACCAAGTTTGTCTCCATGCTGGGCGCTCTTCTTGTGATGTTCTGG 660
QY 661 CCAGTTTCACTGCTCTGCTGAGAGTCCAAATCGAAAGGAAGTCCGCTTCAAC 720
Db 661 CCAGTTTCACTGCTCTGCTGAGAGTCCAAATCGAAAGGAAGTCCGCTTCAAC 720
QY 721 ACCTCTATGCTGTAGCAGTCAAGCTGTGTGACAGCACTTCAGGGTATCTTGGCTCAC 780
Db 721 ACCTCTATGCTGTAGCAGTCAAGCTGTGTGACAGCACTTCAGGGTATCTTGGCTCAC 780
QY 781 CCCCAAGGAAGATCAGCAAGCTTATGTGACAGTCCGGTGTGCGAGAGCGTGGCT 840
Db 781 CCCCAAGGAAGATCAGCAAGCTTATGTGACAGTCCGGTGTGCGAGAGCGTGGCT 840
QY 841 GTGGGTACTGTGTACCTGATCCCTTCTCCGTGCTTCCATGATGCTGTGGCTTGTG 900
Db 841 GTGGGTACTGTGTACCTGATCCCTTCTCCGTGCTTCCATGATGCTGTGGCTTGTG 900
QY 901 GGTGGGCTGATCTCCGTCCGGGGAGCCAACTGCTCGGAGGCTGTGTAAACCAAGTGTG 960
Db 901 GGTGGGCTGATCTCCGTCCGGGGAGCCAACTGCTCGGAGGCTGTGTAAACCAAGTGTG 960
QY 961 GGGATTCGCCACAGCTCCATCATGAGGCTCAACTTCAAGCTTGGGCTGTGAGAG 1020
Db 961 GGGATTCGCCACAGCTCCATCATGAGGCTCAACTTCAAGCTTGGGCTGTGAGAG 1020
QY 1021 ATCATCTACATTGTGCTGTGTGTGATACCGTCCGAGCCGGCAATGGCAATGTTGGC 1080
Db 1021 ATCATCTACATTGTGCTGTGTGTGATACCGTCCGAGCCGGCAATGGCAATGTTGGC 1080
QY 1081 TTCAGAGTCTCTCTCCAGCATTTGGGGAACTCAGCTTGGCCATGTGATAGCTCTCAGCT 1140
Db 1081 TTCAGAGTCTCTCTCCAGCATTTGGGGAACTCAGCTTGGCCATGTGATAGCTCTCAGCT 1140
QY 1141 GGTCTCTGACAGTGTGCTCTCTAAATCTTAAATATGGAAGCACTCATGAGCTTAA 1200
Db 1141 GGTCTCTGACAGTGTGCTCTCTAAATCTTAAATATGGAAGCACTCATGAGCTTAA 1200

QY 1201 TATTGGATGACCAAGTTTCTGGAAGTTTCTTCATTTGGCTGTGGATT 1251
Db 1201 TATTGGATGACCAAGTTTCTGGAAGTTTCTTCATTTGGCTGTGGATT 1251

RESULT 15
AY449382 1251 bp mRNA linear PRI 03-DEC-2003
LOCUS Homo sapiens rhesus blood group D antigen DLO mRNA, partial cds.
DEFINITION AY449382
ACCESSION AY449382
VERSION AY449382.1 GI:38565512
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1251)
Noizat-Pirenne, P. and Aneau-Pirenne, H.
Direct Submission
Submitted (15-OCT-2003) Blood Bank, EFS, 151 Avenue du Marechal de
Latre de Tassigny, Creteil 94010, France
Location/Qualifiers
1..1251
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..>1251
/note="weak D DLO"
/codon_start=1
/product="rhesus blood group D antigen DLO"
/protein_id="AA024079.1"
/db_xref="GI:38565513"
/translation="MSKYPKRSVSRCLPLMALTLAEALILLPYFTTHYDASLEDOGL
VASVOGDLTWAALIGRLTSSFRHSSSVAFNLFMALGVMAIILDLSPRP
SGKRVITLPSIRLATMSALSVLISVDVLRKNDLQLVNVLVSTALGNLRVTSIT
FNDYFHNMMHIIIPFAVFGLSVAVCLPKLPSTGEBDKDQATVPLSLANGLPLMI
FNPSPNSALRSPRIERNNAVFNYYAVASVVAISGSSLAHQKISKTYVASVLA
GVAVGTLCHLIPSPMLAMVLGLVAGLISVGAKYLPGCCNRVLGIPHSIMGNFSL
GLGLGRIYVLVLDTVGAGNGMIGQVLSIGLSIALIVIALTSGLLTGLLNKI
MKRPHAKYDDQVPKPRHLAVGP"

ORIGIN
Query Match 99.5%; Score 1247.8; DB 9; Length 1251;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1249; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTTAAGTACCCGCGGTCTGTCCGGCGTCCCTCTGGCCCTTAACACTG 60
Db 1 ATGAGCTTAAGTACCCGCGGTCTGTCCGGCGTCCCTCTGGCCCTTAACACTG 60

QY 61 GAAGCAGCTCTAATCTCTCTATTTTAAACCACTATGACGCTTCTTAGAGAT 120
Db 61 GAAGCAGCTCTAATCTCTCTATTTTAAACCACTATGACGCTTCTTAGAGAT 120

QY 61 GAAGCAGCTCTAATCTCTCTATTTTAAACCACTATGACGCTTCTTAGAGAT 120
Db 61 GAAGCAGCTCTAATCTCTCTATTTTAAACCACTATGACGCTTCTTAGAGAT 120

QY 121 CAAAAGGGGCTGTGGCATCTATCAAGTTGGCCAAAGTGAACCGTATGGGCGCAT 180
Db 121 CAAAAGGGGCTGTGGCATCTATCAAGTTGGCCAAAGTGAACCGTATGGGCGCAT 180

QY 121 CAAAAGGGGCTGTGGCATCTATCAAGTTGGCCAAAGTGAACCGTATGGGCGCAT 180
Db 121 CAAAAGGGGCTGTGGCATCTATCAAGTTGGCCAAAGTGAACCGTATGGGCGCAT 180

QY 181 GGGCTGGGCTTCTCACTCGAGTTTCCGAGACACGCTGGAGACAGTGGCCCTCAAC 240
Db 181 GGGCTGGGCTTCTCACTCGAGTTTCCGAGACACGCTGGAGACAGTGGCCCTCAAC 240

QY 241 CTCTTCATGTGGGCTTGTGGTGGAGTGGGCAATCTGTGAGACGGCTTCTTAGAGCAG 300
Db 241 CTCTTCATGTGGGCTTGTGGTGGAGTGGGCAATCTGTGAGACGGCTTCTTAGAGCAG 300

QY 301 TTCCCTTCTGGGAAGGTGTCATCACTGTTCAAGTATTCGGCTGGCCACATGAGTGCT 360
Db 301 TTCCCTTCTGGGAAGGTGTCATCACTGTTCAAGTATTCGGCTGGCCACATGAGTGCT 360

QY 361 TTGTCGGTGTGATCTAGTGAATGCTGTCTTGGGGAAGGTCAACTTGGCGCAGTTGGTG 420
Db 361 TTGTCGGTGTGATCTAGTGAATGCTGTCTTGGGGAAGGTCAACTTGGCGCAGTTGGTG 420

QY 421 GTGATGATGTGGTGGAGTGAAGCTTTAGGCAACCTGAGATGTCATCAGTAATATC 480
Db 421 GTGATGATGTGGTGGAGTGAAGCTTTAGGCAACCTGAGATGTCATCAGTAATATC 480

QY 481 TTCAACACAGACTACCACTGAACATGATGCAATCTAGCTTTCCGACCTATTTGGG 540
Db 481 TTCAACACAGACTACCACTGAACATGATGCAATCTAGCTTTCCGACCTATTTGGG 540

QY 541 CTGCTGTGGCCGTGGTCCGCAAGCCCTACCCGAGGGAACGAGATTAAGATCAG 600
Db 541 CTGCTGTGGCCGTGGTCCGCAAGCCCTACCCGAGGGAACGAGATTAAGATCAG 600

QY 601 ACAGCAACGATACCCAGTTTGTCTGCATGCTGGGCGCTCTTCTTGTGGATGTTCTGG 660
Db 601 ACAGCAACGATACCCAGTTTGTCTGCATGCTGGGCGCTCTTCTTGTGGATGTTCTGG 660

QY 661 CCAAGTTTCAACTCTGCTGTGTAAGAGTCCAATGGAAGAAAGATCCGTTTCAAC 720
Db 661 CCAAGTTTCAACTCTGCTGTGTAAGAGTCCAATGGAAGAAAGATCCGTTTCAAC 720

QY 721 ACCTACTAGCTGATGAGTCAAGGCTGAGAGCCATCCAGGTCATCTTGGCTCAC 780
Db 721 ACCTACTAGCTGATGAGTCAAGGCTGAGAGCCATCCAGGTCATCTTGGCTCAC 780

QY 781 CCCCAAGGGAAGATCAGCAAGACTTAATGTCACAGTGCCTGTTGGCAGAGGCGTGGCT 840
Db 781 CCCCAAGGGAAGATCAGCAAGACTTAATGTCACAGTGCCTGTTGGCAGAGGCGTGGCT 840

QY 841 GTGGGTAATCTGTGTGCACTGATCCCTTCCGTGGTCCATGCTGGGTCTTGTG 900
Db 841 GTGGGTAATCTGTGTGCACTGATCCCTTCCGTGGTCCATGCTGGGTCTTGTG 900

QY 901 GCTGGGCTGATCTCCGTCGGGGGAGCCAAATACCTGCCGGGGTCTTGAACCAAGTCTG 960
Db 901 GCTGGGCTGATCTCCGTCGGGGGAGCCAAATACCTGCCGGGGTCTTGAACCAAGTCTG 960

QY 961 GGGATTCGCCACAGCTTCATCATGGGCTAACATCTTCAAGTTGGCTGTGGAGAG 1020
Db 961 GGGATTCGCCACAGCTTCATCATGGGCTAACATCTTCAAGTTGGCTGTGGAGAG 1020

QY 1021 ATCATCTAATGTCGTGCTGTGTGATATCCGTGGAGCCGGAATGSCATGATTTGGC 1080
Db 1021 ATCATCTAATGTCGTGCTGTGTGATATCCGTGGAGCCGGAATGSCATGATTTGGC 1080

QY 1081 TTCCAGGTCTCTCAAGATTTGGGGAATCAGCTTGGCCATCGTGATAGCTTCAAGTCT 1140
Db 1081 TTCCAGGTCTCTCAAGATTTGGGGAATCAGCTTGGCCATCGTGATAGCTTCAAGTCT 1140

QY 1141 GGTCTCCTGACAGTTTGTCTTAAATCTTAAATATGAAAGCACTCATGAGGCTAAA 1200
Db 1141 GGTCTCCTGACAGTTTGTCTTAAATCTTAAATATGAAAGCACTCATGAGGCTAAA 1200

QY 1201 TATTGGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1251
Db 1201 TATTGGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1251

Search completed: March 23, 2005, 02:39:44
Job time : 5464 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2005, 23:12:42 ; Search time 700 Seconds
(without alignments)
10604.794 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1254
Sequence: 1 atgagctcaagtcaccgcg.....attgcgtctgattttaa 1254

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the chance being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1254	100.0	1254	2	AAx86522
2	1252.4	99.9	1354	12	ADN88526
3	1252.4	99.9	1354	12	ADQ47195
4	1252.4	99.9	2811	10	ADQ47195
5	1037	82.7	1810	13	ADQ47195
6	468	37.3	642	12	ADP28536
7	428.8	34.2	627	12	ADP28536
8	365.8	29.2	487	9	ACH43482
9	177	14.1	1805	6	ABK12333
10	175.4	14.0	1792	6	ABK12333
11	175.4	14.0	1792	12	ADM43998
12	165.4	13.2	1351	8	ACD20392
13	165.4	13.2	1853	3	AAc64286
14	165.4	13.2	1949	5	AAH25763
15	165.4	13.2	1952	6	ABK49215
16	165.4	13.2	1952	12	ADN04909
17	165.4	13.2	1952	13	ADN04909
18	165.4	13.2	1968	8	ABX76257
19	165.4	13.2	1970	3	AAc76344
20	159.4	12.7	1679	6	ABK12334

C	21	152.4	12.2	3706	4	AAH18728	Aah18728 Human cDN
C	22	150.2	12.0	2098	6	ABK49216	Abk49216 cDNA enco
C	23	148.8	11.9	2415	10	ADB62576	Adb62576 Human cDN
C	24	147.8	11.8	339	10	ADQ07624	Adq07624 Novel cod
C	25	139.8	11.1	486	4	AAI11615	Aai11615 Probe #15
C	26	139.8	11.1	486	4	ABA53307	Abas3307 Human fce
C	27	139.8	11.1	486	4	AAI32911	Aai32911 Probe #15
C	28	139.8	11.1	486	4	ABA42886	Abas42886 Human bre
C	29	139.8	11.1	486	4	ABA23081	Abas23081 Probe #15
C	30	139.8	11.1	486	4	AAK27012	Aak27012 Human bon
C	31	139.8	11.1	486	4	AAK15568	Aak15568 Human bon
C	32	139.8	11.1	486	4	ABK26596	Abk26596 Human liv
C	33	139.8	11.1	486	5	AAI01542	Aai01542 Probe #15
C	34	139.8	11.1	486	6	ABK01595	Abk01595 Human gen
C	35	135	10.8	1188	10	ADQ07444	Adq07444 Novel cod
C	36	127.2	10.1	954	12	ADR20082	Adr20082 Human imm
C	37	120	9.6	1571	6	ABL90445	Abi90445 Human pol
C	38	113.8	9.1	1840	4	ABL02133	Abi02133 Drosophili
C	39	111.8	8.9	123	4	AAI20829	Aai20829 Probe #10
C	40	111.8	8.9	123	4	ABA65900	Abas65900 Human fce
C	41	111.8	8.9	123	4	AAI46064	Aai46064 Probe #14
C	42	111.8	8.9	123	4	ABA48015	Abas48015 Human bre
C	43	111.8	8.9	123	4	ABA32986	Abas32986 Probe #11
C	44	111.8	8.9	123	4	AAK40045	Aak40045 Human bon
C	45	111.8	8.9	123	4	AAK14315	Aak14315 Human bra

ALIGNMENTS

RESULT 1
AAx86522
ID AAX86522 standard; cDNA; 1254 BP.

XX AAX86522;

AC 04-OCT-1999 (first entry)

XX cDNA sequence of the prevalent allele of the Rhd gene.

XX Allele; Rhesus D antigen; Rhd; weak D phenotype; blood transfusion; ss.

XX Homo sapiens.

OS Homo sapiens.

FH Key Location/Qualifiers
CDS 1..1254
FT /*tag= a

XX WO9937763-A2.

XX 29-JUL-1999.

PF 18-DEC-1998; 98WO-EP008319.

XX 23-JAN-1998; 98EP-00101203.

PA (DRKB-) DRK BLUTSPENDEDIENST BADEN WUERTTEMBERG.

XX Flegel WA, Wagner FF;

XX WPI; 1999-469127/39.

DR P-PDSB; AAY24056.

PT Nucleic acid sequences correlated with Rhesus weak D phenotype. useful for screening blood from donors and recipients for transfusion methods.

XX Disclosure; Fig 2; 64dp; English.

XX The present sequence represents the prevalent allele of the Rhesus D (Rhd) antigen gene. The specification describes a Rhd contributing to or indicative of the weak D phenotype, where the Rhd polynucleotide carries at least one missense mutation as compared to the wild-type Rhd, in its transmembrane and/or intracellular regions, especially in amino acid

CC positions 2-16, 114-149, 179-225 or/and 267-397, with the proviso that
 CC the D antigen does not carry a single missense mutation leading to a
 CC R233V or T263I substitution. The probes and antibodies are useful in the
 CC methods for detection of weak D phenotypes. Red blood cells, from
 CC probands, are useful for the assessment of the affinity, avidity and/or
 CC reactivity of monoclonal anti-D antibodies, polyclonal anti-D antisera or
 CC of anti-globulin or anti-human-globulin antisera. Detecting the presence
 CC of the Rhd associated with weak D phenotype is useful for determining
 CC that a patient in need of a blood transfusion is to be transfused with
 CC RhD negative blood from a donor. Alternatively, testing for weak D
 CC phenotype Rhd in the blood of a donor is useful for determining whether
 CC the donor blood should be excluded for transfusion to patients having
 CC wild type Rhd or weak D types, other than that of the donor weak D type
 CC XX
 SQ Sequence 1254 BP; 241 A; 322 C; 351 G; 340 T; 0 U; 0 Other;

Query Match 100.0%; Score 1254; DB 2; Length 1254;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCTTAAGTACCCGCGGTCTGTCCGGCGTCCCTCTGTGGCCCTTAACATG 60
 Db 1 ATGAGCTTAAGTACCCGCGGTCTGTCCGGCGTCCCTCTGTGGCCCTTAACATG 60

QY 61 GAAGCAGCTCATCT 120
 Db 61 GAAGCAGCTCATCT 120

QY 121 CAAAAGGGGCTGTGGCAGTCCATCAAGTTGGCCAAATCGATGAGCGGCGCAT 180
 Db 121 CAAAAGGGGCTGTGGCAGTCCATCAAGTTGGCCAAATCGATGAGCGGCGCAT 180

QY 181 GGCCTTGGGCTTCTCACTCGAGTTTCCGAGACACAGCTGAGCAGTGTGCTTCAAC 240
 Db 181 GGCCTTGGGCTTCTCACTCGAGTTTCCGAGACACAGCTGAGCAGTGTGCTTCAAC 240

QY 241 CTCTTCAAGTGTGCGCTTGTGTGTCAGGCGCAATCTGTGTCAGCGCTTCTGAAGCAG 300
 Db 241 CTCTTCAAGTGTGCGCTTGTGTGTCAGGCGCAATCTGTGTCAGCGCTTCTGAAGCAG 300

QY 301 TTCCCTTCTGGGAAGGTGTGATCACATGTTCAATTCGAGTTCGCGGCAATGAGTGTCT 360
 Db 301 TTCCCTTCTGGGAAGGTGTGATCACATGTTCAATTCGAGTTCGCGGCAATGAGTGTCT 360

QY 361 TTGTCGGTGTGATCTCAGTGAATGCTGTTGGGGAAGTCACTTGGGCGAGTTGTG 420
 Db 361 TTGTCGGTGTGATCTCAGTGAATGCTGTTGGGGAAGTCACTTGGGCGAGTTGTG 420

QY 421 GTGATGTGTGTGTGAGTGTGACAGCTTTAGCACTTGAAGATGTCACTGATATATC 480
 Db 421 GTGATGTGTGTGTGAGTGTGACAGCTTTAGCACTTGAAGATGTCACTGATATATC 480

QY 481 TTCAACACAGATACCAATGAAATGATGACATCTACGTTGTGACGCTATTTTGGG 540
 Db 481 TTCAACACAGATACCAATGAAATGATGACATCTACGTTGTGACGCTATTTTGGG 540

QY 541 CTGTCTGTGGCTGTGTGCTGCAAGCCTTCAACCGAAGGAAACGAGATAAAGATAG 600
 Db 541 CTGTCTGTGGCTGTGTGCTGCAAGCCTTCAACCGAAGGAAACGAGATAAAGATAG 600

QY 601 ACGGCAACATACCAATGTTGTGTGCAATGCTGTGGGCGCTTCTTGTGATGTCTGG 660
 Db 601 ACGGCAACATACCAATGTTGTGTGCAATGCTGTGGGCGCTTCTTGTGATGTCTGG 660

QY 661 CCAAGTTTCACTCTGCTGCTGGAAGTCCAATCGAAGAAAGAAATGCGGTGTCAAC 720
 Db 661 CCAAGTTTCACTCTGCTGCTGGAAGTCCAATCGAAGAAAGAAATGCGGTGTCAAC 720

QY 721 ACTTACTATGTGTGACAGTCAAGCGTGTGACAGCATCTCAAGGTCATCTTGGCTAC 780
 Db 721 ACTTACTATGTGTGACAGTCAAGCGTGTGACAGCATCTCAAGGTCATCTTGGCTAC 780

QY 781 CCCCAAGGGAATACGACAAAGCTTATGTGACAGTGTGCGGTGTGGCAGAGCGGTGCT 840
 Db 781 CCCCAAGGGAATACGACAAAGCTTATGTGACAGTGTGCGGTGTGGCAGAGCGGTGCT 840

Db 781 |||||
 CCCCAGGGAAGATACAGAAAGCTTATGTGACAGTGTGCGGTGTGGCAGAGCGGTGCT 840

QY 841 GTGGGTACCTGTGTGACAGTCAATCCCTTCCGCGGTGGCTTCCATGTGCTGGGTCTTGG 900
 Db 841 GTGGGTACCTGTGTGACAGTCAATCCCTTCCGCGGTGGCTTCCATGTGCTGGGTCTTGG 900

QY 901 GCTGGGTGATCTTCCGCGGTGGGAGGCAAGTACCTGCGGAGGTGTGTAAACGAGTCTG 960
 Db 901 GCTGGGTGATCTTCCGCGGTGGGAGGCAAGTACCTGCGGAGGTGTGTAAACGAGTCTG 960

QY 961 GGGATTCGCCAGAGCTCCATCATAGGCTTCACTTACGTTGTGCTGGGTCTTGGAGAG 1020
 Db 961 GGGATTCGCCAGAGCTCCATCATAGGCTTCACTTACGTTGTGCTGGGTCTTGGAGAG 1020

QY 1021 ATCATCTACATTTGTGCTGTGCTGTGTGATATACCGTGTGAGCGGCAATGAGTATG 1080
 Db 1021 ATCATCTACATTTGTGCTGTGCTGTGTGATATACCGTGTGAGCGGCAATGAGTATG 1080

QY 1081 TTCCAGTCTCTCTCTCAAGATTGGGGAATCAAGCTTGGCCATGTGATAGCTTCAAGTCT 1140
 Db 1081 TTCCAGTCTCTCTCTCAAGATTGGGGAATCAAGCTTGGCCATGTGATAGCTTCAAGTCT 1140

QY 1141 GGTCTCTGACAGGTTTGTCTCTTAATCTTAATATGAAAGCACTTCAAGAGCTTAA 1200
 Db 1141 GGTCTCTGACAGGTTTGTCTCTTAATCTTAATATGAAAGCACTTCAAGAGCTTAA 1200

QY 1201 TATTTTGAATGACCAAGTTTCTGAAAGTTTCTCATTTGGCTGTGTGATTTTAA 1254
 Db 1201 TATTTTGAATGACCAAGTTTCTGAAAGTTTCTCATTTGGCTGTGTGATTTTAA 1254

RESULT 2
 ADN88526
 ID ADN88526 standard; cDNA, 1354 BP.
 XX
 AC ADN88526;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE Human RHD gene cDNA sequence.
 XX
 KW mutation detection; exon nine; human; RHD; gene; Rh negative blood type;
 XX ss.
 OS Homo sapiens.
 XX
 PN JF2004089057-A.
 XX
 PD 25-MAR-2004.
 XX
 PF 30-AUG-2002; 2002JP-00254009.
 XX
 PR 30-AUG-2002; 2002JP-00254009.
 XX
 PA (NISE-) NIPPON SEKIYU GISHA.
 XX
 PA (WAKT) WAKUNAGA SEIYAKU KK.
 XX
 DR WPI; 2004-262639/25.
 XX
 PT Nucleic acid molecule useful for determining Rh negative blood type, has
 PT fragment which hybridizes to polymucleotide having mutation of guanine
 PT nucleotide to another nucleotide in 3' terminal of ninth exon.
 XX
 PS Claim 2; SEQ ID NO 1; 16pp; Japanese.
 XX
 CC The invention comprises a nucleic acid that can detect a mutation of a
 CC guanine residue to another nucleotide in the 3' terminal region of the
 CC ninth exon of the human RHD gene. The nucleic acid of the invention is
 CC useful for determining Rh negative blood type before birth and
 CC immediately after birth. The present cDNA sequence represents the human
 CC RHD gene.
 XX

SQ Sequence 1354 BP; 278 A; 346 C; 369 G; 361 T; 0 U; 0 Other;
 Query Match 99.9%; Score 1252.4; DB 12; Length 1354;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGGCGCTGCGCCCTCTGGGCCCTTAACACTG 60
DB 1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGGCGCTGCGCCCTCTGGGCCCTTAACACTG 60
QY 61 GAAGCAGCTCTCATCTCTCTCTCTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
DB 61 GAAGCAGCTCTCATCTCTCTCTCTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
QY 121 CAAAGAGGCGCTGCGGCGCATCTATCAAGTTGGCCCAAGATCTGACCGTATGGCGGCATT 180
DB 121 CAAAGAGGCGCTGCGGCGCATCTATCAAGTTGGCCCAAGATCTGACCGTATGGCGGCATT 180
QY 181 GGCCTTGGGCTTCTCTACCTCGAGTTTCCGAGACACACAGCTGAGCAGTGTGCGCTTCAC 240
DB 181 GGCCTTGGGCTTCTCTACCTCGAGTTTCCGAGACACACAGCTGAGCAGTGTGCGCTTCAC 240
QY 241 CTCTTCACTGCTGGCGCTTGTGTGCAATGCGGCAATCTCTGCTGACGCGCTTCTGAGCAG 300
DB 241 CTCTTCACTGCTGGCGCTTGTGTGCAATGCGGCAATCTCTGCTGACGCGCTTCTGAGCAG 300
QY 301 TTCCCTTCTGGGAGAGGTGTCTATCAGTCTGTTGAGTGGGCGCACCATGAGTGTCT 360
DB 301 TTCCCTTCTGGGAGAGGTGTCTATCAGTCTGTTGAGTGGGCGCACCATGAGTGTCT 360
QY 361 TTGTGCGTGTGATCTCTCAGTGTGTCTGTCTTGGGAGAGTCACTTGGCGCAGTTGGTG 420
DB 361 TTGTGCGTGTGATCTCTCAGTGTGTCTGTCTTGGGAGAGTCACTTGGCGCAGTTGGTG 420
QY 421 GTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
DB 421 GTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
QY 481 TTCAACACAGTACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 481 TTCAACACAGTACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 541 CTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
DB 541 CTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY 601 ACAGCAACGATACCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
DB 601 ACAGCAACGATACCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
QY 661 CCAAGTTTCACTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
DB 661 CCAAGTTTCACTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
QY 721 ACCTACTATCTGTAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
DB 721 ACCTACTATCTGTAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 781 CCCCAGGAGGATACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
DB 781 CCCCAGGAGGATACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
QY 841 GTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
DB 841 GTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
QY 901 GCTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
DB 901 GCTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
QY 961 GGGATTTCCCAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
DB 961 GGGATTTCCCAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
  
```

QY 1021 ATCATCTACATTTGT 1080
 DB 1021 ATCATCTACATTTGT 1080
 QY 1081 TTCAGGTCTCTCTCAGCATTTGGGGAATCTCAGCTTGGCCATTCGTATAGCTCTCAGCTCT 1140
 DB 1081 TTCAGGTCTCTCTCAGCATTTGGGGAATCTCAGCTTGGCCATTCGTATAGCTCTCAGCTCT 1140
 QY 1141 GGTCTCTGACAGGTTTGTCTCTAAATCTTAAATATGAAAGACCTCATGAGGCTTAA 1200
 DB 1141 GGTCTCTGACAGGTTTGTCTCTAAATCTTAAATATGAAAGACCTCATGAGGCTTAA 1200
 QY 1201 TATTTGATGACCAAGTTTCTGAAATTTCTCATTTGGCTGTGATTTTAA 1254
 DB 1201 TATTTGATGACCAAGTTTCTGAAATTTCTCATTTGGCTGTGATTTTAA 1254

RESULT 3
 ADO47195
 ID ADO47195 standard; cDNA; 1354 BP.
 AC ADO47195;
 AC 15-JUL-2004 (first entry)
 DT
 DE cDNA sequence of human Rhd blood group antigen.
 DE immunoglobulin; B cell; germline; region V; region D; region J;
 KW recombination-promoting protein; recombination activating gene 1; RAG-1;
 KW RAG-2; ss; terminal deoxynucleotidyltransferase; Tdt; V(D)J recombinase;
 KW Rhd blood group antigen.
 OS Homo sapiens.
 XX
 XX WO2004029249-A1.
 XX
 XX 08-APR-2004.
 XX
 XX 30-SEP-2003; 2003WO-AU001286.
 XX
 XX 30-SEP-2002; 2002US-0415024P.
 XX
 XX (CENT-) CENTENARY INST CANCER MEDICINE & CELL BI.
 XX
 XX Jolly C;
 XX
 XX WPI; 2004-316126/29.
 XX
 XX
 XX New vector comprising one or more immunoglobulin regions selected from V,
 PT D, and J regions, a 5' flanking region, and a 3' flanking region, useful
 PT for reverting cell lines to a pro-B cell-like state or to a germline-like
 PT state.
 PT
 XX
 XX
 XX Example 4; Page 322-323; 382pp; English.
 XX
 XX The specification describes a method for the preparation of
 CC immunoglobulins. The method uses a vector for reverting cell lines to a
 CC pro-B cell-like state or to a germline-like state, by replacing one or
 CC more of the immunoglobulin regions V, D, and J of the rearranged
 CC immunoglobulin gene with the germ-line or unrearranged V, D or J regions.
 CC The vector can comprise a polynucleotide sequence encoding a
 CC recombination-promoting protein, selected from recombination activating
 CC gene 1 (RAG-1), RAG-2, terminal deoxynucleotidyltransferase (Tdt). These
 CC proteins collectively constitute a V(D)J recombinase. The method is
 CC useful for the preparation of immunoglobulin heavy and/or light chains
 CC and their fragments. The present sequence encodes a human Rhd blood group
 CC antigen, and is in the course of the invention.
 CC
 XX
 SQ Sequence 1354 BP; 278 A; 346 C; 369 G; 361 T; 0 U; 0 Other;
 Query Match 99.9%; Score 1252.4; DB 12; Length 1354;
 Best Local Similarity 99.9%; Pred. No. 0;

Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAGCTCTAAGTACCAGGAGTCTGTCGGGCGTGCCTCCCTTGGGCGCTTAACATCG 60
Db 1 ATGAGCTCTAAGTACCAGGAGTCTGTCGGGCGTGCCTCCCTTGGGCGCTTAACATCG 60
QY 61 GAAGCAGCTCTCATTTCTCTCTTCTATTTTACCACAGTACGCTTCTTAAGAGAT 120
Db 61 GAAGCAGCTCTCATTTCTCTCTTCTATTTTACCACAGTACGCTTCTTAAGAGAT 120
QY 121 CAAAAGGGGCTGTGGCATCTTCAAGTTGGGCAAGATCTGACCGTGATGCGGCCATT 180
Db 121 CAAAAGGGGCTGTGGCATCTTCAAGTTGGGCAAGATCTGACCGTGATGCGGCCATT 180
QY 181 GGGTTGGGCTTCTCTCACTCGAGTTCCGGAGACACAGCTGAGAGAGTGGGCTTCAAC 240
Db 181 GGGTTGGGCTTCTCTCACTCGAGTTCCGGAGACACAGCTGAGAGAGTGGGCTTCAAC 240
QY 241 CTCTTCATGCTGGCGGTTGGTGTGCAGTGGCAATCTGCTGAGACGGCTTCTGAGCCAG 300
Db 241 CTCTTCATGCTGGCGGTTGGTGTGCAGTGGCAATCTGCTGAGACGGCTTCTGAGCCAG 300
QY 301 TTCCCTTCTGGAGAGTGTGTCATCAACAGTTCACTATTCGGCTGGCCACCATGAGTCT 360
Db 301 TTCCCTTCTGGAGAGTGTGTCATCAACAGTTCACTATTCGGCTGGCCACCATGAGTCT 360
QY 361 TTGTGCTGCTGATCTCAGTGTGATCTTGGGGAAGTCACTTGGCGCAGTGGTG 420
Db 361 TTGTGCTGCTGATCTCAGTGTGATCTTGGGGAAGTCACTTGGCGCAGTGGTG 420
QY 421 GTGATGCTGTGGTGAAGTGAAGCTTAGGCAACCTGAGAGTGTATGTAATATC 480
Db 421 GTGATGCTGTGGTGAAGTGAAGCTTAGGCAACCTGAGAGTGTATGTAATATC 480
QY 481 TTCAACACAGACTACCATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 481 TTCAACACAGACTACCATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 541 CTGCTGTGGGCTGTGGTGTGCTGCAAGGCTTACCCGAGGAGCGAGATAAAGATCAG 600
Db 541 CTGCTGTGGGCTGTGGTGTGCTGCAAGGCTTACCCGAGGAGCGAGATAAAGATCAG 600
QY 601 ACAGCAACGATACCAAGTTGTGTCAGTGTGGGCGCCCTTCTTGTGTGATGTTCTGG 660
Db 601 ACAGCAACGATACCAAGTTGTGTCAGTGTGGGCGCCCTTCTTGTGTGATGTTCTGG 660
QY 661 CCAAGTTTCAACTTGTCTGTGCTGAGAAATTCGAATGCAAGAAAGAAATGCGGTTCAC 720
Db 661 CCAAGTTTCAACTTGTCTGTGCTGAGAAATTCGAATGCAAGAAAGAAATGCGGTTCAC 720
QY 721 ACCCTATGCTGTGAGAGTCAAGGTGTGAGCAAGCACTCAGAGTCACTTGGCTCAC 780
Db 721 ACCCTATGCTGTGAGAGTCAAGGTGTGAGCAAGCACTCAGAGTCACTTGGCTCAC 780
QY 781 CCCCAGGAGAAATCAGCAAGACTTATGTGCAAGTGTGTCAGAGAGCGGTGCT 840
Db 781 CCCCAGGAGAAATCAGCAAGACTTATGTGCAAGTGTGTCAGAGAGCGGTGCT 840
QY 841 GTGGGTACCTCGGTGACCTGATCCCTTCTCCGTGCTTCCATGCTGTGGTCTTGTG 900
Db 841 GTGGGTACCTCGGTGACCTGATCCCTTCTCCGTGCTTCCATGCTGTGGTCTTGTG 900
QY 901 GCTGGGCTGATCTCGCTGGGAGGCAAGTACCTGCGGGGTTGTTAAACGAATGCTG 960
Db 901 GCTGGGCTGATCTCGCTGGGAGGCAAGTACCTGCGGGGTTGTTAAACGAATGCTG 960
QY 961 GGGATTTCCCAACAGCTCCATCATGGGCTTAACTTCAAGTTGCTGGTCTGTGAGAG 1020
Db 961 GGGATTTCCCAACAGCTCCATCATGGGCTTAACTTCAAGTTGCTGGTCTGTGAGAG 1020
QY 1021 ATCATCTATGATGTGCTGTGAGCTGATGATGATGATGATGATGATGATGATGATG 1080
Db 1021 ATCATCTATGATGTGCTGTGAGCTGATGATGATGATGATGATGATGATGATGATG 1080

```

```

QY 1081 TTCCAGTCTCTCTCAGCATTTGGGAACTCAGCTTGGCCATCGATAGCTCAGCTCT 1140
Db 1081 TTCCAGTCTCTCTCAGCATTTGGGAACTCAGCTTGGCCATCGATAGCTCAGCTCT 1140
QY 1141 GGTCTCTGACAGGTTTGTCTCTTAAATCTTAAATATGAAAGACCTCATGAGCTTAA 1200
Db 1141 GGTCTCTGACAGGTTTGTCTCTTAAATCTTAAATATGAAAGACCTCATGAGCTTAA 1200
QY 1201 TATTTGATACCAAGTTTCTGAAAGTTCCCATTTGGCTGTGATTTTAA 1254
Db 1201 TATTTGATACCAAGTTTCTGAAAGTTCCCATTTGGCTGTGATTTTAA 1254

```

RESULT 4
ADE09785
ADE09785 standard; DNA; 2811 BP.

AC ADE09785;

DT 29-JAN-2004 (first entry)

DE Novel DNA-related contig nucleotide sequence #507.

KW novel gene; novel protein; tissue marker; molecular weight marker;
KW chromosome marker; genetic disorder; contig; ds.

OS Unidentified.

PN MO2003054152-A2.

PD 03-JUL-2003.

PF 10-DEC-2002; 2002MO-US039555.

PR 10-DEC-2001; 2001US-0339733P.

PR 11-DEC-2001; 2001US-0339453P.

PR 14-MAR-2002; 2002US-0365091P.

PR 14-MAR-2002; 2002US-0365384P.

PR 12-APR-2002; 2002US-0372381P.

PR 22-APR-2002; 2002US-0372615P.

PR 22-APR-2002; 2002US-00128558.

PR 24-APR-2002; 2002US-0376045P.

XX (HYSB-) HYSBQ INC.

XX Tang YF, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AU, Wehman T, Weng G, Zhou P, Drmanac RT, Wang Z;
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX WPI; 2003-569235/53.

XX New polynucleotide, useful for expressing recombinant proteins for
PT analysis, characterization or therapeutic use, or as markers for tissues
PT in which the corresponding protein is preferentially expressed.

XX Disclosure; SEQ ID NO 2329; 1177bp; English.

XX The invention comprises the amino acid and coding sequences of novel
CC proteins. The DNA and protein sequences of the invention are useful as:
CC markers for tissues in which the corresponding protein is preferentially
CC expressed; as molecular weight markers on gels; as chromosome markers or
CC tags; to identify chromosomes or to map related gene positions; and to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. The present DNA sequence was used in the
CC exemplification of the invention.

XX Sequence 2811 BP; 658 A; 690 C; 659 G; 804 T; 0 U; 0 Other;

SO Query Match 99.9%; Score 1252.4; DB 10; Length 2811;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Query Match 82.7%; Score 1037; DB 13; Length 1810;
Best Local Similarity 93.9%; Pred. No. 2,7e-289;
Matches 1079; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

```
QY 106 GCTTCCTTAAAGATCAAAAGGGGCTGCGGCACTCTATCAAGTTGGCCAAAGATCGACC 165
DB 447 GCTTCCTTAAAGATCAAAAGGGGCTGCGGCACTCTATCAAGTTGGCCAAAGATCGACC 506
QY 166 GTGATGGCGGCAATGGGCTTGGGCTTCTCACTCGAGTTCCGGAGACAGACGCTGAGC 225
DB 507 GTGATGGCGGCAATGGGCTTGGGCTTCTCACTCGAGTTCCGGAGACAGACGCTGAGC 566
QY 226 AGTGTGGCTTCAACTCTTCACTGCTGGCGCTTGGTGCAGTGGGCAATCTGCTGAGC 285
DB 567 AGTGTGGCTTCAACTCTTCACTGCTGGCGCTTGGTGCAGTGGGCAATCTGCTGAGC 626
QY 286 GGGTCTGAGGCAAGTTCCTTCTGGGAGGTGTCATACCTGTTGAGTATTCGAGCTG 345
DB 627 GGGTCTGAGGCAAGTTCCTTCTGGGAGGTGTCATACCTGTTGAGTATTCGAGCTG 686
QY 346 GCGACCATGAGTCTTGTGTGGTGTGATCTCAGTGTGATGCTGTTGGGGAGGTCAAC 405
DB 687 GCGACCATGAGTCTTGTGTGGTGTGATCTCAGTGTGATGCTGTTGGGGAGGTCAAC 746
QY 406 TTGGGCGAGTGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 465
DB 747 TTGGGCGAGTGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 806
QY 466 GTCATCAGTAATATCTTCAACACAGACTACCATGATGACATGATGACATGATGATG 525
DB 807 GTCATCAGTAATATCTTCAACACAGACTACCATGATGACATGATGACATGATGATG 866
QY 526 GCAAGCTATTTTGGGCTGTCTGTGGCTGTGTGGCTGTGTGGCTGTGTGGCTGTGTGG 585
DB 867 GCAAGCTATTTTGGGCTGTCTGTGGCTGTGTGGCTGTGTGGCTGTGTGGCTGTGTGG 926
QY 586 GAGGATTAAGATCAAGACAGACATACCACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCT 645
DB 927 GAGGATTAAGATCAAGACAGACATACCACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCT 986
QY 646 TTGTGATGTTCTGGCCCAAGTTTCACTGTCTGTGTGAGAAATCCAAATCGAAAGAG 705
DB 987 TTGTGATGTTCTGGCCCAAGTTTCACTGTCTGTGTGAGAAATCCAAATCGAAAGAG 1046
QY 706 AATCCCGGTTCAACACTATGCTGTAGCACTGACGCTGTGTGACAGCCATCTCAAGG 765
DB 1047 AATCCCGGTTCAACACTATGCTGTAGCACTGACGCTGTGTGACAGCCATCTCAAGG 1106
QY 766 TCATCTTGGCTCACTCCCAAGGAGATCAGCAAGACTTATGTGACAGTGGGGGTG 825
DB 1107 TCATCTTGGCTCACTCCCAAGGAGATCAGCAAGACTTATGTGACAGTGGGGGTG 1166
QY 826 GCGAGGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 885
DB 1167 GCGAGGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1226
QY 886 GTGTGGGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 945
DB 1227 GTGTGGGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1286
QY 946 TGTAAACGAGTGTGGGATTTCCCAAGCTCATGATGAGGCTCACTCACTTCACTGCTG 1005
DB 1287 TGTAAACGAGTGTGGGATTTCCCAAGCTCATGATGAGGCTCACTCACTTCACTGCTG 1346
QY 1006 GGTCTCTTGAAGATCACTTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1065
DB 1347 GGTCTCTTGAAGATCACTTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1406
QY 1066 AATGCGATATTTGCTTCAAGGCTCTCTCAAGCATTTGGGAACTCAGCTGGCCATGCTG 1125
DB 1407 AATGCGATATTTGCTTCAAGGCTCTCTCAAGCATTTGGGAACTCAGCTGGCCATGCTG 1466
```

```
QY 1126 ATAGCTCAAGCTGTGTCTCTGACAGGTTTGCTCTTAATCTTAATATGAAAGCA 1185
DB 1467 ATAGCTCAAGCTGTGTCTCTGACAGGTTTGCTCTTAATCTTAATATGAAAGCA 1526
QY 1186 CCTGATGAGGCTAAATTTTGTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGT 1245
DB 1527 CCTGATGAGGCTAAATTTTGTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGT 1586
QY 1246 GGATTTTAA 1254
DB 1587 GGATTTTAA 1595

RESULT 6
ADP28536
ID ADP28536 standard; DNA; 642 BP.
XX
AC ADP28536;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein encoding sequence SEQ ID #534.
XX
KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; de; human secreted protein.
XX
OS Homo sapiens.
XX
PN M02004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003M0-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
```


PR 29-AUG-2002; 2002US-040666P.
 PR 17-SEP-2002; 2002US-0410946P.
 PR 17-SEP-2002; 2002US-0410947P.
 PR 17-SEP-2002; 2002US-0410948P.
 PR 17-SEP-2002; 2002US-0410949P.
 PR 17-SEP-2002; 2002US-0410953P.
 PR 17-SEP-2002; 2002US-0410957P.
 PR 17-SEP-2002; 2002US-0410958P.
 PR 17-SEP-2002; 2002US-0410959P.
 PR 17-SEP-2002; 2002US-0410960P.
 PR 17-SEP-2002; 2002US-0410961P.
 PR 17-SEP-2002; 2002US-0411019P.
 PR 17-SEP-2002; 2002US-0411022P.
 PR 17-SEP-2002; 2002US-0411023P.
 PR 17-SEP-2002; 2002US-0411024P.
 PR 17-SEP-2002; 2002US-0411032P.
 PR 17-SEP-2002; 2002US-0411035P.
 PR 17-SEP-2002; 2002US-0411037P.
 PR 17-SEP-2002; 2002US-0411041P.
 PR 17-SEP-2002; 2002US-0411045P.
 PR 17-SEP-2002; 2002US-0411046P.
 PR 17-SEP-2002; 2002US-0411048P.
 PR 17-SEP-2002; 2002US-0411052P.
 PR 17-SEP-2002; 2002US-0411055P.
 PR 17-SEP-2002; 2002US-0411073P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411012P.
 PR 17-SEP-2002; 2002US-0411111P.
 PR 18-APR-2003; 2003US-0463700P.
 PR 18-APR-2003; 2003US-0463708P.
 PR 18-APR-2003; 2003US-0463716P.
 PR 02-MAY-2003; 2003US-0463732P.
 PR 02-MAY-2003; 2003US-0467199P.
 PR 02-MAY-2003; 2003US-0467201P.
 PR 02-MAY-2003; 2003US-0467203P.
 PR 19-MAY-2003; 2003US-0467230P.
 PR 19-MAY-2003; 2003US-0471306P.
 PR 22-MAY-2003; 2003US-0472420P.
 PR 09-JUN-2003; 2003US-0472430P.
 PR 09-JUN-2003; 2003US-0476609P.
 PR 08-JUL-2003; 2003US-0476641P.
 PR 08-JUL-2003; 2003US-0485218P.
 PR 08-JUL-2003; 2003US-0485223P.
 PR 14-JUL-2003; 2003US-0485224P.
 PR 14-JUL-2003; 2003US-0485325P.
 PR 14-JUL-2003; 2003US-0486446P.
 PR 15-JUL-2003; 2003US-0486480P.
 PR 15-JUL-2003; 2003US-0486891P.
 PR 08-AUG-2003; 2003US-0486960P.
 PR 08-AUG-2003; 2003US-0493341P.
 PR 08-AUG-2003; 2003US-0493370P.
 PR 08-AUG-2003; 2003US-0493573P.
 PR 08-AUG-2003; 2003US-0493577P.
 XX
 XX
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
 XX
 PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
 PI Halebbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
 XX
 XX
 DR WPI; 2004-348438/32.
 XX
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 XX genetic, bacterial and viral diseases.
 XX
 PS Claim 1; SEQ ID NO 509; 428bp; English.
 CC
 CC The present invention relates to an isolated nucleic acid molecule
 CC encoding a polypeptide which is believed to be cytotoxic,
 CC antiinflammatory, immunosuppressive, antibacterial and antiviral. The

CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein encoding sequence. The
CC present sequence is available on WIPower and is not in the specification
XX
Sequence 627 BP; 119 A; 171 C; 180 G; 157 T; 0 U; 0 Other.

Query Match	34.28	Score 428.8	DB 12	Length 627
Best Local Similarity	77.68	Pred. No. 2,4e-113		
Matches 605	Conservative 0	Mismatches 22	Indels 153	Gaps 1
QY	154	CAGATGTGACCGTGTATGAGGCGCCATTGGCTTGGGCTTCCTACCTCCAGTTTCCGGAGA	213	
Db	1	CAAGATGTACCGTGTATGAGGCGCCATTGGCTTGGGCTTCCTACCTCCAGTTTCCGGAGA	60	
QY	214	CACAGCTGAGACGTGTGCTTCAACCTTTATGCTGGCGCTTGTGTGACAGTGGACA	273	
Db	61	CACAGCTGAGACGTGTGCTTCAACCTTTATGCTGGCGCTTGTGTGACAGTGGACA	120	
QY	274	ATCCTGTGAGACGGCTTCTCTGAGCCAGTTCCTTTGGGAAAGTGTATCACACTGTC	333	
Db	121	ATCCTGTGAGACGGCTTCTCTGAGCCAGTTCCTTTGGGAAAGTGTATCACACTGTC	180	
QY	334	AGTATTCGCGTGGCCACCATGAGTCTTGTGTGCTGATCTCAGTGGATGCTGTCTG	393	
Db	181	AG-----	182	
QY	394	GGGAAGTCACTTGGCCAGTGTGTGTATGCTGCTGTGGAGTGCACCTTAAAGC	453	
Db	183	-----	182	
QY	454	AACTGAGAGTGCATCATGTAATATTCTTCAACA CAGACTACCATGAACATGATGAC	513	
Db	183	-----	207	
QY	514	ATCTACGTGTTGCGACGCTATTTTGGGCTGTCTGTGGCCGTGGTCCCAAGCCCTTA	573	
Db	208	TTCTACGTGTTGCGACGCTATTTTGGGCTGTCTGTGGCCGTGGTCCCAAGCCCTTA	267	
QY	574	CCCGAGGGAA CGAGAGTAAGATCAGACAGCAACGATACCCAGTTTGTCTGCCATGCTG	633	
Db	268	CCCAAGGGAACGAGAGTAAGATCAGACAGCAACGATACCCAGTTTGTCTGCCATGCTG	327	
QY	634	GGCGCCTCTTCTTGTGTGATGTTCTGCGCAAGTTTCAACTCTGCTGCTGAGAAATGCA	693	
Db	328	GGCGCCTCTTCTTGTGTGATGTTCTGCGCAAGTTTCAACTCTGCTGCTGAGAAATGCA	387	
QY	694	ATCGAAGGAAGATGCGCTTTCACACTACTATGCTGTAGACGTACAGCTGTGTACA	753	
Db	388	ATCGAAGGAAGATGCGCTTTCACACTACTATGCTGTAGACGTGTAGACGTGTGTACA	447	
QY	754	GCCATCTCAGGATCATCTGTGCTCAACCCCAAGGGAGATCGCAAGACTTATGTGCAC	813	
Db	448	GCCATCTCAGGATCATCTGTGCTCAACCCCAAGGGAGATCGCAAGACTTATGTGCAC	507	
QY	814	AGTCCGCTGTTTGGCAGAGGCGTGTGTGTACTCTGTGTACCTGATTCCTCTTCCG	873	
Db	508	AGTCCGCTGTTTGGCAGAGGCGTGTGTGTACTCTGTGTACCTGATTCCTCTTCCG	567	
QY	874	TGGCTTTCATGAGTGGGCTTGTGGCTGGGGCTGATCCCTGGTGGGGAACCAAGTAC	933	
Db	568	TGGCTTTCATGAGTGGGCTTGTGGCTGGGGCTGATCCCTGGTGGGGAACCAAGTAC	627	

DE	Human foetal brain cDNA #4207.
XX	
KM	Human; 88; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW	genome mapping; biodiversity; genetic disorder.
XX	
OS	Homo sapiens.
XX	
PN	US2003073623-A1.
PD	
XX	
PF	30-JUL-2001; 2001US-00918995.
XX	
PR	30-JUL-2001; 2001US-00918995.
XX	
PA	(DRMA/) DRMANAC R T.
PA	(LABA/) LABAT I.
PA	(STAC/) STACHE-CRAIN B.
PA	(DICK/) DICKSON M C.
PA	(JONE/) JONES L W.
PI	Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX	
DR	WPI; 2003-615964/58.
XX	
PT	New polynucleotide sequences obtained from various cDNA libraries, useful
PT	as hybridization probes, as oligomers for PCR, for chromosome and gene
PT	mapping, in the recombinant production of protein, or in generating
PT	antisense DNA or RNA.
PS	
XX	Claim 1; SEQ ID NO 30694; 44pp; English.
XX	
CC	The invention relates to an isolated polynucleotide comprising any one of
CC	38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC	determined by the technique of SBH (sequencing by hybridisation). Also
CC	included is a purified polypeptide comprising a sequence corresponding to
CC	a reading frame of the novel polynucleotide. The nucleic acid sequences
CC	are useful in diagnostics as expressed sequence tags (EST) for
CC	identifying expressed genes or for physical mapping of the human genome,
CC	in forensics, in assessing biodiversities, or in identifying mutations
CC	responsible for genetic disorders and other traits. The nucleotide
CC	sequences are also useful as hybridisation probes, as oligomers for PCR,
CC	for chromosome and gene mapping, in the recombinant production of
CC	protein, or in generating antisense DNA or RNA. The purified polypeptide
CC	is useful for generating antibodies specific for it. The present sequence
CC	is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC	for this patent did not form part of the printed specification, but was
CC	obtained in electronic format directly from USPTO at
CC	seqdata.uspto.gov/sequence.html?docID=20030073623
XX	
XX	Sequence 487 BP; 86 A; 120 C; 156 G; 119 T; 0 U; 6 Other;
XX	
Query	March 29.2%; Score 365.8; DB 9; Length 487;
Beet	Local Similarity 99.5%; Pred. No. 3.8e-95;
Matches	367; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	706 AATGCCGCTGTTCACACCACTTACTATGCTTAGCAGTCACGCTGGTGAACGCAATCTCAGAG 765
DB	56 ATTCCCGTGTTCAACACCTACTATGCTGTGAAGATCAAGCTGGTGTACAGCATCTCAGG 115
QY	766 TCATCTCTTGAGCTCACCCCAAGAGGAGATCAGCAAGACTTATGTGCACAGTGGGTGTTG 825
DB	116 TCATCTCTTGAGCTCACCCCAAGAGGAGATCAGCAAGACTTATGTGCACAGTGGGTGTTG 175
QY	826 GCAGAGGCGGTGGCTGGTGGTATACCTCGGTGTACCTGATCCCTTCGCGGTGGTGCATG 885
DB	176 GCAGAGGCGGTGGCTGGTGGTATACCTCGGTGTACCTGATCCCTTCGCGGTGGTGCATG 235
QY	886 GTGCTGGCTCTTGTGGCTGGCTGTATCTCCGTGGGGAGCCAAATGACTGCGGGGTGT 945
DB	236 GTGCTGGCTCTTGTGGCTGGCTGTATCTCCGTGGGGAGCCAAATGACTGCGGGGTGT 295
QY	946 TGTAAACGAGTGTGGGATTTCCCAACAGCTTCATCATGAGGTACCACTTCACTTGGT 1005

Db	296	TGTATCCGAGTGGCTGGGGGATTCGCCAGAGCTCCATCATGGGCTACAACTTCAGCTTGCTG	355
Oy	1006	GGTCTGCTTGGAGAGATCATCTACATTGCTGCTGCTGCTTGTATACCGTCGAGCCGGC	1065
Db	356	GGTCTGCTTGGAGAGATCATCTACATTGCTGCTGCTGCTTGTATACCGTCGAGCCGGC	415
Oy	1066	AATGGCATG 1074	
Db	416	AATGGCATG 424	
RESULT 9			
ABK12333	1D	ABK12333 standard; cDNA; 1805 BP.	
XX	AC	ABK12333;	
XX	DT	05-JUN-2002 (first entry)	
XX	DE	cDNA encoding human nonerythroid Rh glycoprotein RhBG.	
XX	XX	Human; nonerythroid Rh glycoprotein; Rhbg; RhBG; antibody; immunogen;	
KW	KW	Rh type B transporter gene; polypeptic transporter-type protein;	
KX	KX	ion transporter; chromosome 1q21.3; gene; ss.	
OS	XX	Homo sapiens.	
FH	XX	Key	location/Qualifiers
FT	FT	CDS	39..1415
FT	FT	misc_feature	1248..1377
FT	FT	polyA_signal	1769..1774
FT	FT		/*tag= C
FT	FT		/standard_name= "PolyA signal"
FT	FT		/note= "Atypical polyadenylation site"
XX	XX	W0200216396-A2.	
PD	XX	28-FEB-2002.	
PF	XX	17-AUG-2001; 2001WO-US025881.	
PR	XX	21-AUG-2000; 2000US-0226767P.	
PA	XX	(NYBL-) NEW YORK BLOOD CENT INC.	
P1	XX	Huang C, Liu Z;	
DR	XX	WPI; 2002-280905/32.	
DR	XX	P-PSDB; MAU78091.	
XX	XX	Homologs of mouse or human nonerythroid Rh glycoproteins, Rhbg or RhBG,	
XX	XX	respectively, useful in the production of antibodies which are useful for	
XX	XX	detecting Rhbg or RhBG glycoproteins in a sample.	
XX	XX	Claim 1; Fig 1a; 59pp; English.	
XX	XX	The present invention relates to a new protein or peptide comprising an	
XX	XX	amino acid sequence having at least 60% identity to a sequence comprising	
XX	XX	455 (mouse nonerythroid Rh glycoprotein homologue (Rhbg)) or 458 (human	
XX	XX	nonerythroid Rh glycoprotein homologue (RhBG)) amino acids, fully defined	
XX	XX	in the specification. The antibody of the invention is useful for	
XX	XX	detecting an Rhbg or an RhBG glycoprotein in a sample, by contacting the	
XX	XX	sample with antibody under conditions suitable for binding, assessing the	
XX	XX	specific binding to the antibody, and thus detecting the presence of an	
XX	XX	epitope of Rhbg or RhBG in the sample. The nucleic acids of the invention	
XX	XX	are useful as probes for detecting transporter genes and particularly Rh	
XX	XX	type B transporter genes including e.g. NH ₄ ⁺ ion transporters. The	
XX	XX	present nucleic acid sequence is that of the human RhBG gene located on	

CC chromosome 1q21.3. This sequence encodes the human RhBG protein of the
CC invention. RhBG is a polytypic transporter-type protein

XX Sequence 1805 BP; 351 A; 550 C; 515 G; 389 T; 0 U; 0 Other;

Query Match 14.1%; Score 177; DB 6; Length 1805;
Best Local Similarity 49.5%; Pred. No. 3.8e-40;
Matches 456; Conservative 0; Mismatches 465; Indels 0; Gaps 0;

```

QY 141 CTATCAAGTTGGCCAGATCTGACCGGATGAGGGGCGGCTTGGCTTCTCTCACTTC 200
DB 218 CTACCCAGCTTCAGAGACGTGCATGCCATGCTTCTTGCGGCTTGGCTTCTCTATGCT 277
QY 201 GAATTTCCGAGACACAGCTGAGAGAGTGGCTTCAACCTTTTCATGCTGGCGCTTGG 260
DB 278 CTTCCTGACAGCGTTTACGGCTTTCAGAGCGTGGGCTTCACTTCTCTCGCGCTTGGC 337
QY 261 TGTGACGGGGAATCTCTGCTGAGCGCTTCTGAGCCGTTCCCTCTTGGAGAGTGTGT 320
DB 338 CCGCAGTGTGTCACTGCTGTCAGGGCTTCTTCCACTCTTCCACGGTGGCCACATCCA 397
QY 321 CATACACTGTTTCAGTATTCGCTGCGCCACATGAGTGTGCTTGTGCTGATCTCACT 380
DB 398 TGTGCGCTGTGAGAGATGATCAATGCTGACTTTTGTGGGGGCGCTGCTCATCTCTT 457
QY 381 GGATGCTGCTTGGGGAAGTCACTTGGCGGAGTTGGTGTGATGATGCTGATGAGAGT 440
DB 458 TGTGCTCGCTTGGGGAACCGGGCCACACCGCTGCTGCTCATGCGCTTGTGAGAGT 517
QY 441 GACAGCTTTAGGCAACTGAGAGATGATCATGTAATATCTTCAACACAGCTACCAT 500
DB 518 GGTGCTGTTTGCATCATGATGATTTGTGCTCTTCTATCTCTCGGGGGTGAAGATGCGG 577
QY 501 GAACTATGATGACATCTACGTGTTGCAAGCTTATTTTGGGCTGTCTGTGGCTGATGCT 560
DB 578 AGGCTCCATGACTATCCACACTTTGGTGGCTTCTGCGCTCGCTTGTGGGGTCTT 637
QY 561 GCCAAGCCTCTACCCGAGGGAAGAGAGATTAAGATCAGACACACATACCACTTT 620
DB 638 GTACAGGCGCCCAAGCTGAGGAAGAGACACCGCCAGGGCTCCCTTACCATTCAGACT 697
QY 621 GTCTGCACTGCTGGCGGCTCTTCTTGTGATGATTTGCGCAAGTTTCAACTGCTGCT 680
DB 698 CTTCGCTTATGATGAGACATCTTCTGTGATCTTCTGCGCTTACCTTCAATGCTGACT 757
QY 681 GCTGAGAGTCCATTCAGGAAGAGATGCTGTTTCAACCTTATCTGTATGACAT 740
DB 758 CACAGCGCTGGGGGCTGGGAGCATGCGACGCGCCCTCACAACATACTACTCTGCTGC 817
QY 741 CAGGCTGTGACACGCACTCTCAGGGTATCTTGGCTCACCCCAAGGGAAGATCAGCA 800
DB 818 CAGCACTTGGACCTTTGCTTGTGACGCTTGTAGGGGAAGATGGAGGCTTACAT 877
QY 801 GACTTATGTCACAGTCCGCTGTTGGCAGAGCGTGGCTGTGGATACCTCTGTCACCT 860
DB 878 GTTCCATCTCAAAATGACAGCGCTGAGGGGCTTGTGGGAGACCTCAAGTAAAT 937
QY 861 GATCCCTTCTCCGCTGCTGCAATGAGTGGCTGTGGCTGGCTGCTATCTCCGCTCG 920
DB 938 GATGCTGACACCTTTGGGGCTCTGGCAGCTGAGCTTCTTGGCTGGGACGTCTCCACGCT 997
QY 921 GGAAGCAAGTACCTGCGGGGCTGTGTGAACGAGTGTGGGATTTCCCAAGCTCCAT 980
DB 998 GGGGTACAAATTTTACGCGCCATCTCTGAATCAAAATTCAAAGTCAAGACATGTGG 1057
QY 981 CATGGCTACCAATTCAGCTTGTGGGTCTGCTGGAAGATCATCTATGTTGCTGCT 1040
DB 1058 AGTCCCAACATCTCATGAGATGCGGGGCTCTCGGGGCGCTTCTGAGGCTTGTGGC 1117
QY 1041 GGTGCTTGAATACCGTGGAGC 1061
DB 1118 TGGACTTGGCACCCATGGAAGC 1138

```

RESULT 10

AB211480 standard; cDNA; 1792 BP.

AB211480;

20-JAN-2003 (first entry)

Human polynucleotide SEQ ID NO 362.

Human: genome mapping; gene therapy; food supplement; virus; fungus;
cell-proliferative disorder; neurodegenerative disease; bacterial;
Parkinson's disease; Alzheimer's disease; autoimmune disease;
multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
haemostatic; vulnery; fungicide; immunosuppressive; dermatological;
antiarthritic; gene; ss.

Homo sapiens.

WO200270539-A2.

12-SEP-2002.

05-MAR-2002; 2002WO-US005095.

05-MAR-2001; 2001US-00799451.

(HYSE-) HYSEQ INC.

Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
P I Weinman T, Wang J, Wang D, Drmanac RT;

WPI; 2002-759812/82.

P-PSDB; ABP69263.

New polynucleotides comprising sequences assembled from expressed
sequence tags (ESTs), useful for treating cell-proliferative,
neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
or coagulation disorders.

Claim 1; SEQ ID NO 362; 1012bp + Sequence Listing; English.

The invention relates to an isolated polynucleotide (I) comprising a
nucleotide sequence selected from any of 948 sequences (AB21119-
AB212065) or their mature protein coding portion, active domain coding
protein or complementary sequences. The polynucleotides are useful for
identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, wound, burn, lymphoid disorders,
CC platelet or coagulation disorders, myeloid or lymphoid disorders,
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
XX directly from WIP0 at ftp.wipo.int/pub/published_pcl_sequences

Sequence 1792 BP; 340 A; 547 C; 511 G; 394 T; 0 U; 0 Other;

Query Match 14.0%; Score 175.4; DB 6; Length 1792;
Best Local Similarity 49.4%; Pred. No. 1.1e-39;
Matches 455; Conservative 0; Mismatches 466; Indels 0; Gaps 0;

```

QY 141 CTATCAAGTTGGCCAGATCTGACCGGATGAGGGGCGGCTTGGCTTCTCTCACTTC 200
DB 211 CTACCCAGCTTCAGAGACGTGCATGCCATGCTTCTTGCGGCTTGGCTTCTCTATGCT 270

```


QY	201	GAGTTTCCGGAAACACAGCTGGAGCAGTGTGGCTTCAACCTCTTCATGCTGGCGCTTGG	260
Db	271	CTTCTCGACGCTTACGGCTTTAGCAGCGTGGCTTACCTTCTCTTGGCGGCTTTGC	330
QY	261	TGTGCAGTGGGCAATCTGCTGGACGGCTTCTGTAGCCAGTTCCCTTCTGGAGAGTGTG	320
Db	331	CCTGCAGTGTGTCCACACTGTGTCCAGGGGCTTTCTCCACTCCTTCCACAGGTGGCAATCCA	390
QY	321	CATCACACTGTTCAGTATTGGGCTGGCCACCAATGAGTGTCTTTGTGGTGTGATCTCAGT	380
Db	391	TGTTGGCGTGGAGACATGATCAATGTCTACTTTTGTGGGGGGCGGTGCTCATCTCTT	450
QY	381	GGATGCTCTCTTGGGGGAAGCTCAACTTGGCGGAGTTGTGTGTGATGTGTCTGTGGAGGT	440
Db	451	TGTGTCCGCTCGGGCAAAACCGGGCTTACCCAGCTGTGCTCATGGCCCTGCTGGAGGT	510
QY	441	GACACCTTTAGGCAACCTGAGGATGGTCAATCAGTAATATCTTCAACACAGACTACACAT	500
Db	511	GGTGTGTTTGGCAATCAATGATTTGTGTCTCTTCAATCTCTCTGGGGGTAGAGATGCCGG	570
QY	501	GAACATGATGACATCTAAGTGTTCGACGCTTAATTTTGGGCTGTCTGTGGCCCTGTGCT	560
Db	571	AGGCTCACTGACTATTCACACCTTTGTGTCTACTTCCGGGCTGTCTCTTTGGGGTGTCT	630
QY	561	GCCAAAGCTCTTACCCCGAGGGAGCGAGGATTAAGATCAGACAGCAAGATACCAATT	620
Db	631	GTACAGGGCCCCAGCTGGAGAAAGCAGACGCCCGGCGCTCCGTCTACATTCAGACCT	690
QY	621	GTCTGCATGCTGAGGGGCGCCCTCTTCTGTGTGAGATGTTCTGGGCCAAGTTTCAACTGTCT	680
Db	691	CTTGCCCATGATTTGGGACCATCTTCTGTGTGATCTTCTGGCTTACGCTTCAATGTGTCACT	750
QY	681	GCTGAAAGTCCATTCGAAAGAGAGATGCCGTGTTCAACACTTACTATCTGTAGCAGT	740
Db	751	CACAGCGCTGGGGGCGTGGCAGACATCCGAGCGGCCCTCAACACTACTACTCCCTGGCTGC	810
QY	741	CAGCTGTGTACAGCCATCTCAGGGTCACTCTTGTGCTACCCCCAAGGAGATCAGCA	800
Db	811	CAGCACCTTGGCACCCTTGTGCTTGTCAAGCCCTTGAAGGGAGAAATGGGAGGCTTGACAT	870
QY	801	GACTTATGTGCACACTGTGCGGTGTTGGCAGAGAGCGGTGTGTGGGTAACCTCGTGTCACT	860
Db	871	GGTCCACATCCAAATGTACAGCGCTGTGGAGAGGGGTGTGTGTGGGACCTCAAGTGAAT	930
QY	861	GATCCCTTCTCCGTGGCTTGCATGATGTGTGGTCTTGTGGCTGGGCTGATCTCCGTGG	920
Db	931	GATGTGACACCCCTTTGGGGGCTGTGGCACTGGCTTCTTGGCTGGGACGTGTCTCCACGCT	990
QY	921	GGGAGCCAAAGTACTGTGCCGGGGTGTGTGAACCGAGTGTGGGGAATTCGCCACAGCTCAT	980
Db	991	GGGGTACAAAGTTCTTACGGCCATCTTGAATCAAAATTCAAAGTCCAAAGACATGTGG	1050
QY	981	CATGGGCTTACCACTTCAGCTTGTGTGTGTGCTGTTGGAGAGATCATCTACATTGTGTGCT	1040
Db	1051	AGTCCACAACCTCCATGTGAGATGCCGGGGGTCTGTGGGGGCGCTTCGTGGGGGTCTTGTGGC	1110
QY	1041	GATGCTGATACCGTGGAGC 1061	
Db	1111	TGGACTTCCACCCATGAAGC 1131	
RESULT 11			
ADM43998			
ID ADM43998 standard: cDNA, 1792 BP.			
XX ADM43998;			
XX 03-JUN-2004 (first entry)			
XX Novel human arginine-rich protein cDNA #362.			
XX 8a; gene; human; arginine-rich protein; cancer; inflammation;			
XX genetic disorder.			

XX	Homo sapiens.
OS	
XX	
XX	US2004053250-A1.
FN	
XX	
PD	18-MAR-2004.
XX	
XX	21-NOV-2002; 2002US-00302172.
PF	
XX	05-MAR-2001; 2001US-00799451.
PR	05-MAR-2002; 2002MO-US005095.
PR	20-AUG-2002; 2002US-00225251.
XX	
PA	(TANG/) TANG Y T.
PA	(XUEA/) XUE A.
PA	(DRMA/) DRMANAC R T.
XX	
PI	Tang YT, Xue A, Drmanac RT;
XX	
DR	WPI; 2004-238579/22.
PT	
PT	New isolated arginine-rich protein-like polynucleotides and polypeptides,
PT	useful for diagnosing and/or treating conditions associated with aberrant
PT	activity of the arginine-rich polypeptides, such as cancer and
PT	inflammation.
XX	
PS	Disclosure; SEQ ID NO 362; 51pp; English.
XX	
CC	The invention relates to an isolated polynucleotide. The methods and
CC	compositions of the present invention are useful for the diagnosis and/or
CC	treatment of diseases or conditions associated with aberrant expression
CC	or activity of the arginine-rich protein-like polypeptides, such as
CC	cancer and inflammation. They can also be used in forensics, gene
CC	mapping, identification of mutations responsible for genetic disorders,
CC	and in assessing biodiversity. The present sequence represents a novel
CC	human arginine-rich protein cDNA.
XX	
SQ	Sequence 1792 BP; 340 A; 547 C; 511 G; 394 T; 0 U; 0 Other;
Query Match	14.0%; Score 175.4; DB 12; Length 1792;
Best Local Similarity	49.4%; Pred. No. 1.1e-39;
Matches 455; Conservative	0; Mismatches 466; Indels 0; Gaps 0
OY	141 CTATCAAGTTGGCCAAAGATCGACCCTGATGGCGGCCATTGGGCTTCCTCACCTC
Dd	211 CTACCCAAGCTTCCAGACGCTGCATGCCATGTCTTCGTGGGCTTGACTTCCTCATGGT
OY	201 GAGTTTCCGGAGACAACAGCTGGAGCAGTGTGGCTTCAACCTCTTCATGCTGGCGCTTGS
Dd	271 CTTCCCTGAGGGTTAACGGCTTCAGCAGCGTGGGCTTCACTTCCTCGCCGCGCTTTCG
OY	261 TGTCAGTGGGCAATCTCTGCTGAGACGGCTTCTTGAGCCAGTTCCCTTCGCGGAAGGTGGT
Dd	331 CCTCAGATGGTCCACACATCGTGCACGGGCTTCTCCACTCTTCCACCGGTGGCCACATCCA
OY	321 CATCAACCTGTTCAAGTATTTGGGCTGGCCACCATGATGTGTTGTGCGTGTGATCTCAGT
Dd	391 TGTGGCGGTGAGAGCATGATCATGTCTGACTTTTGTGTGGGGGCGGTGCTCATCTCTT
OY	381 GGATGCTGTCTTGGGGGAAGTCAACTTGGCGCAGTTGGTGTGATGTGTGCTGCTGAGAGGT
Dd	451 TGGTCCGCTCTGGGGCAAGACCGGGCCTAACCAAGCTGTGCTCATGTGCCCTGTGSAAGT
OY	441 GACAGCTTTAGGCAACCTGAGAGTGTATCATCACTAATCTTCAACACAGACTACCAAT
Dd	511 GGTCTGTGTTGGCATCATGATGATTTGTCTCTTCATCTCTCCGGGGGTGAGATGCCGG
OY	501 GAACATGATGACATCTACGATGTTGCGACGCTAATTTGGGCGTGTGTTGGCGCTGGAGCT
Dd	571 AGGCTTCATGACTATCCACACCTTTGGTGTCTTCTGAGGCTGTCTCTTTTCCGGGTTCT
OY	561 GCCAAGCTCTTACCCGAGGGAACGAGATTAAGATCAGACAGCAACGATACCAAGTTT
	620

Db 631 GTACAGACCCAGCTGAGAGAGCAAGCACCGCGGCTCCGTACATTCAGACCT 690
QY 621 GTCTGCATCTGAGGAGCCCTCTTCTTGTGAGATGTTGAGCCAGTTTCACTTCTCT 680
Db 691 CTTCGCGATGATGGAGCACTTCTCTGTGATCTTGTGCTTAAGTTAAAGCTGCACT 750
QY 681 GCTGAGAAAGTCCATGAAAGAAAGATGCGGTTCACACCTACTATGCTGTAGCACT 740
Db 751 CACAGCGCTGGGGGCTGGGAGCATGAGCGGCCCTCAACATCTATCTCCCTGGCTGC 810
QY 741 CACGCTGTGAGACGATCTCAGGCTCATCTTGGCTCACCCCAAGGAGATACGAA 800
Db 811 CAGACCCCTTGGAGCATCTTGGCTCTTGTACCCCTTGAAGGAGATGGAGGCTTGACAT 870
QY 801 GACTTATGTGACAGATGCGGCTGTGAGAGGCGGCTGTGGGATCTGCTGACCT 860
Db 871 GGTCCACATCTCAAAATGACGCGCTGCTGAGAGGGGTGTGGTGGGACCTCAAGTGAAT 930
QY 861 GATCCCTTCTCGTGGCTTGCATGTGCTGGGTCTTGTGGCTGGGCTGATCTCCGTCG 920
Db 931 GATGCTGACACCCCTTGGGGCTCTGCGACGCTGCTTGTGGCTGGGACGTCTCCAGCT 990
QY 921 GGGAGCCAGTACTCTGCGCGGGGTGTGTAAACGAGGCTGGGATTTCCCAAGTCCAT 980
Db 991 GGGGTAAAGATTCTTACGCGCCCATCTTGAATCAAAATTCAAATGCAAGACATGTGG 1050
QY 981 CATGGGCTTCAACTTCACTTGTCTGCTGGGTCTGCTTGTGAGAGATCATCTTATGTGCT 1040
Db 1051 AGTCCAAACCTTCCATGAGATGCGGGGGTCTGTGGGCGCTCTCGGGGGTCTTGTGGC 1110
QY 1041 GGTGCTTGAATACGCTGGAGC 1061
Db 1111 TGGACTTGCACCCATGAAGC 1131
RESULT 12
ACD20392/C
ID ACD20392 standard; DNA; 1351 BP.
XX
AC ACD20392;
XX
DT 26-AUG-2003 (first entry)
XX
DE DNA encoding human NOV15 protein.
XX
KW Human; NOX; inflammatory disorder; demyelination disease; stroke;
KW renal disorder; infection; cardiomyopathy; atherosclerosis; acne;
KW hyperextension; pancreatitis; Von Hippel-Lindau; endometriosis; fertility;
KW scleroderma; cirrhosis; inflammatory bowel disease; Crohn's disease;
KW haemophilia; autoimmune disease; allergy; AIDS;
KW graft versus host disease; Alzheimer's disease; arthritis; pain;
KW Parkinson's disease; Huntington's disease; obesity; diabetes;
KW hair growth; hair loss; asthma; schizophrenia; glomerulonephritis;
KW lupus erythematosus; psoriasis; antidiabetic; anorectic; metabolic;
KW nocturnal; neuroprotective; cytosolic; antibacterial; viral; virologic;
KW protozoa; antiatherosclerotic; hypotensive; cerebroprotective;
KW antiinflammatory; gynaecological; antifertility; dermatological;
KW hepatocellular; haemostatic; immunosuppressive; antiallergic;
KW anticholinergic; anticommutant; antiseborrhoeic; antiaschemic;
KW neuroleptic; anti-HIV; analgesic; nephrotoxic; antipsychotic; gene; de.
XX
OS Homo sapiens.
XX
PN WO200298917-A2.
XX
PD 12-DEC-2002.
XX
PF 12-FEB-2002; 2002WO-US022049.
XX
PR 12-FEB-2001; 2001US-0268221P.
PR 13-FEB-2001; 2001US-026846P.
PR 14-FEB-2001; 2001US-0268646P.
PR 14-FEB-2001; 2001US-0268665P.

PR 15-FEB-2001; 2001US-0269136P.
PR 16-FEB-2001; 2001US-0269310P.
PR 16-FEB-2001; 2001US-0269530P.
PR 15-MAR-2001; 2001US-0276405P.
PR 16-MAR-2001; 2001US-0276399P.
PR 16-MAR-2001; 2001US-0276703P.
PR 23-MAR-2001; 2001US-0278199P.
PR 28-MAR-2001; 2001US-0279274P.
PR 30-MAR-2001; 2001US-0280238P.
PR 02-APR-2001; 2001US-0280899P.
PR 08-AUG-2001; 2001US-0310797P.
PR 14-AUG-2001; 2001US-0312284P.
PR 14-SEP-2001; 2001US-0322284P.
PR 14-SEP-2001; 2001US-0322285P.
PR 18-OCT-2001; 2001US-0330293P.
PR 31-OCT-2001; 2001US-0335104P.
PR 31-OCT-2001; 2001US-0335109P.
PR 21-NOV-2001; 2001US-0332127P.
PR 28-NOV-2001; 2001US-0331772P.

(CURA-) CURAGEN CORP.

Guo X, Fernandes E, Li L, Kekuda R, Liu Y, Lette M, Spytek KA;
Ji W, Casman SJ, Boldog FL, Patrajan M, Vernet CM, Ballinger RA,
Malyankar UM, Tcherven VT, Blalock AD, Gusev VY, Rasbell L;
Mezes PD, Ellerman K, Heyes M, Herrmann JL, Shimkets RA, Ioime N;
Pena CE, Shenoy SG, Taupier RJ, Gerlach V, Gorman L;
WPI, 2003-148650/14.
DR P-SDB; ABO15004.
XX

Novel NOX polypeptide useful for identifying an agent that binds to the
PT polypeptide and for treating cardiomyopathy, atherosclerosis,
PT hypertension, infertility, scleroderma, cirrhosis, and inflammatory bowel
PT disease.

Claim 9; Page 115; 566pp; English.

XX
PS The present invention relates to the isolation of novel human
XX polypeptides referred to as NOX (NOV1-NOV37), variants of these
XX proteins, and the polynucleotide sequences encoding them. The NOX
XX proteins of the invention share homology to various types of protein
XX families such as zinc finger-like proteins, enzymes, receptors, and
XX lipoproteins. The sequences of the invention may be useful in the
XX manufacture of a medicament for treating a syndrome associated with a
XX human disease. For example, they can be used to treat inflammatory
XX disorders, demyelination disease, renal disorders, infections,
XX cardiomyopathy, atherosclerosis, hypertension, stroke, pancreatitis, Von
XX Hippel-Lindau, endometriosis, fertility, scleroderma, cirrhosis,
XX inflammatory bowel disease, Crohn's disease, haemophilia, autoimmune
XX diseases, allergies, graft versus host disease, Alzheimer's disease,
XX arthritis, Parkinson's disease, Huntington's disease, obesity, diabetes,
XX acne, hair growth/loss, asthma, schizophrenia, AIDS, pain,
XX glomerulonephritis, lupus erythematosus, and psoriasis. ACD20372-ACD20427
XX represent DNA sequences encoding the NOX polypeptides of the invention.
XX Note: SEQ ID Nos 113-460 are known sequences used for homology purposes
SQ Sequence 1351 BP; 325 A; 372 C; 389 G; 265 T; 0 U; 0 Other;

Query Match 13.2%; Score 165.4; DB 8; Length 1351;
Best Local Similarity 49.1%; Pred. No. 7.5e-37;
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGGCCCAAGATCGACCGGATGGGCGCATTTGGGCTTCTCAACCTC 200
Db 1175 CTACCAAGCTTCCAGAGCGTACGATGCTTCTGTGGCTTCTCTCAATGAC 1116
QY 201 GAGTTTCCGAGACACAGCTGAGCAAGTGTGCTTCAACTTCTCATGCTGGCGCTGG 260
Db 1115 TTTCCTGACAGCGCTACGCGCTTTCAGCGCCGCTTCAACTTCTGTGGAGCTTGG 1056
QY 261 TGTCAAGTGGCAATCTGCTGAGCGCTTCTGAGCCAGTTCCCTTGTGGAAGTGGT 320

Db 1055 CATCCAGTGGGCGCTGCTCATGACGGGCTGGTTCACCTTCTTACAGAACCGCTACATCGT 996
 Qy 321 CATCACACTGTTGAGTATTCGGCTGCGCACCATGAGTCTTTGTCGGTGTGATTCAGT 380
 Db 995 CGTGGGGGTGGAGAACTCATCAACGCTGACTTCTGTGTGGCCCTCTCTCGCGGCTT 936
 Qy 381 GGAATGCTGCTTGGGAGAGGTCAACTTGGCGCAGTGTGGTGTGATGTGCTGTGAGGT 440
 Db 935 TGGGGGAGTTCTGGGTAAAGTCAGCCCATTCAGCTCTCATCATCTTCTTCCAAGT 876
 Qy 441 GACAGCTTTAGGCACTGAGAGTGTTCATCAGTAATATCTTCAACACAGATCAACCAT 500
 Db 875 GACCTCTTGTGCTGTGATGATGATTCCTTAACTCTGTAAAGGTGAAGGTGACAG 816
 Qy 501 GAACATGATGACATCTACGTGTTGCGAGCTTATTTTGGGCTGTGCTGTGCTGTGCT 560
 Db 815 AGGCTCATATACCATTCACACATTTGGCGCTACTTTGGGCTCAAGTGAACCGGATCT 756
 Qy 561 GCCAAGCCTCTACCCGAGGAAACGAGATTAAGATCAGACAGCAAGTACCAGTTT 620
 Db 755 CTACCGACGCACTTAGAGAGAGAGAGAGACAGAAATTTGTGTACCAAGTGGACCT 696
 Qy 621 GTCTGCATGCTGGGCGCCCTCTTCTGTGTGATGTTGTGGCCAAAGTTTCAACTCTGCT 680
 Db 695 CTTTGCATATGATGACCCCTCTTCTGTGTGATGTTGTGGCCCAAGCTTCAACTCAGCAT 636
 Qy 681 GCTGAGAGTTCATCGAAGAAAGAAATGCGCTGTTCACACCTACTATGCTGTGAGAGT 740
 Db 635 ATCTTACCATGGGAGACGACGACGACGCGCATACACCTTACTGCTCTTGGACAC 576
 Qy 741 CACGCTGTGACAGCCATCTCAGGGTCACTCTTGGCTCACCCCAAGGAGATCAGCAA 800
 Db 575 CTGCGTCTTACTCGGTGGCAATATCAGAGCCCTGACACAAAGGCAAGCTGACAT 516
 Qy 801 GACTTATGTGACAGTCCGCTGTGTGAGAGAGCGCTGTGCTGTGCTGTGCTGTGCT 860
 Db 515 GGTGCAATCAGAAATCCACGCTCGCAGAGAGGGGTGGCGGTGGTATCCCTGTGTGAT 456
 Qy 861 GATCCCTTCTCGGCGCTGCAATGCTGTGAGGTCTTGTGGCTGTGGGTGTGATCTCGTGG 920
 Db 455 GATGCTCATGCTTACGCTGCTGCTCATCATGCGCTTGTCTGCGGCAATCTTCACCT 396
 Qy 921 GGGAGCCAGTACCTGCGGGGGGTGTGTAAACGAGTCTGGGGATTTCCCAAGCTCAT 980
 Db 395 GGGTTTGTATACCTGACCCCATTCCTGAGAGTCCCGGCTGCACATTCAGAGACATGTGG 336
 Qy 981 CATGGCTACACTTCACTTGTGCTGTGCTGTGCTGTGAGAGATCATCTTCACTTTGTG 1035
 Db 335 CATTAACAAATCTGCATGTCATTTCTGTCATCATAGCGGAGATCGTGGTGTGTG 281

RESULT 13

AAC64286
 ID AAC64286 standard; cDNA; 1853 BP.

AC AAC64286;

XX 23-FEB-2001 (first entry)

DE Human membrane-associated protein HUMP-13 cDNA.

XX Human membrane-associated protein; HUMAP; transgenic organism;
 KM drug screening; cell signalling modulator; agonist; antagonist;
 KM cell differentiation modulator; cell proliferation modulator;
 KM cell differentiation disorder; cancer; cell differentiation disorder;
 KM developmental disorder; cell signalling disorders; endocrine disorder;
 KM hyperparathyroidism; hyperparathyroidism; infection;
 KM pancreatic disorder; diabetes mellitus; immunological disorder;
 KM hereditary neuropathy; gonadal steroid hormone associated disorder;
 KM infertility; ss.

OS Homo sapiens.

XX

PN W0200065054-A2.
 XX
 PD 02-NOV-2000.
 XX
 PF 20-APR-2000; 2000MO-US010884.
 XX
 PR 23-APR-1999; 99US-0130694P.
 PR 23-JUN-1999; 99US-0140580P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Hillman JL, Bandman O, Tang YT, Lal P, Yue H, Reddy R;
 PI Azimzai Y, Baughn MR;
 XX
 DR WPI; 2000-687346/67.
 DR P-PSDB; AAB29656.
 XX
 PT Human membrane-associated protein, useful for diagnosis and treatment of
 PT cell signalling, cell differentiation and cell proliferation disorders
 PT such as cancer, and for identifying agonists and antagonists.
 XX
 PS Claim 4; Page 96; 99p; English.
 XX
 CC The invention relates to 17 human membrane-associated proteins, HUMAP-1
 CC to HUMAP-17 (AAB29644-B29660) and the cDNAs encoding them (AAC64274-
 CC C64290). The invention also relates to expression constructs, host cells
 CC and transgenic organisms comprising a HUMAP nucleic acid sequence; the
 CC recombinant preparation of a HUMAP; methods of screening compounds for
 CC their ability to modulate HUMAP activity or expression; and
 CC pharmaceutical compositions comprising a HUMAP protein, a HUMAP agonist
 CC or HUMAP antagonist. The HUMAPs acts as modulators of cell signalling,
 CC differentiation and proliferation. A HUMAP is useful for screening a
 CC compound for effectiveness as an agonist or antagonist of HUMAP activity.
 CC The protein, or the identified agonist or antagonist is useful for
 CC treating a disease or condition associated with decreased or increased
 CC expression of functional HUMAP. A HUMAP nucleic acid is useful for
 CC screening a compound for its ability to alter expression of that
 CC particular HUMAP gene. A wide variety of disease may be treated using
 CC compositions of the invention. These diseases include cell proliferative
 CC disorders (e.g., actinic keratosis, arteriosclerosis); cancer (e.g.,
 CC breast, bladder, bone marrow, brain and uterine cancer); cell
 CC differentiation disorders, in particular developmental disorders (e.g.,
 CC renal tubular acidosis, anaemia, Cushing's syndrome, achondroplasia,
 CC epilepsy, and muscular dystrophy); cell signalling disorders, in
 CC particular endocrine disorders such as hypothyroidism and pituitary
 CC disorders resulting from lesions such as thrombosis, disorders associated
 CC with hyperparathyroidism (e.g., acromegaly); disorders associated with
 CC hypothyroidism (e.g., goitre); hyperparathyroidism; pancreatic disorders
 CC such as type I or type II diabetes mellitus; infections; immunological
 CC disorders; hereditary neuropathies (e.g., neurofibromatosis); and
 CC disorders associated with gonadal steroid hormones (e.g., infertility,
 CC endometriosis, polycystic ovary syndrome, osteoporosis, Leydig cell
 CC deficiency and gynecomastia). Antibodies which specifically bind HUMAP
 CC may be used for the diagnosis of disorders associated with the expression
 CC of HUMAP, or in assays to monitor patients being treated with HUMAP or
 CC agonists, antagonists or inhibitors of HUMAP. The present sequence
 CC represents a HUMAP cDNA of the invention
 XX
 SQ Sequence 1853 BP; 367 A; 563 C; 508 G; 415 T; 0 U; 0 Other;
 Query Match 13.2%; Score 165.4; DB 3; Length 1853;
 Best Local Similarity 49.1%; Pred. No. 8.9e-27;
 Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;
 Qy 141 CTATCAAGTTGGCCAAAGATTTGACCGTATGCGGCCCATTTGGGCTTCTTCACCTC 200
 Db 207 CTACCCAGGCTTCAGAGACGTGACGTGATGCTTCTGTGGGCTTCTTCATGAC 266
 Qy 201 GACTTTCGGAGACACAGCTGAGCAGTGTGAGCTTCAACCTTCAATGTGGCGCTTGG 260
 Db 267 TTTCTGCAAGCGCTTACGGCTTACGCCGCTGGCTTCAACTTCTGTGGCAGGCTTGG 326
 Qy 261 TGTGAGTGGGCAATCTGCTGTGACGCGCTTCTGAGCAGCTTCTTGTGGAAAGTGTGT 320

```

Db      327 CATCCAGTGGCGCTGTCATGACAGGAGCTGTTCACCTTCAACAGACCGCTACATCT 386
Qy      321 CATCACACTGTGATGATTCGGCTGGCCACATGAGTGTCTTGGTCTGATCTCACT 380
Db      387 CGTGGGCGTGGAGAACCTCATCAACGCTGACTTGGCGCTCTGTGTGGCTGACCTT 446
Qy      381 GATGCTGTCTTGGGAGAGTCAACTTGGCGAGTGGAGTGGATGATGATGATGATGATGAT 440
Db      447 TGGGGGAGTTCCTGGTAAAGTCAAGCCCATTTAGCGCTCATCATGACTTTCTTCAAGT 506
Qy      441 GACAGCTTAAAGCACTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 500
Db      507 GACCTCTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 566
Qy      501 GAAATGATGACATCTACGCTGTCAGAGCTATTTTGGGCTGTGCTGGCTGGCTGGCTGGCT 560
Db      567 AGGCTCATGACCATTCACACATTTGGCGCTTACTTTGGGCTCAAGTACCGGATCTT 626
Qy      561 GCCAAGCTCTTACCCGAGGGAACGAGAGATTAAGATCAGACGACGATACCCAGTTT 620
Db      627 CTACGAGCGCAACTTAAAGCAGAGCAGAGAGACAGAACTTGTATCAAGTGGACCT 686
Qy      621 GTCTGCATGCTGGGCGCTCTTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 680
Db      687 CTTCGCAATGATGGACACCTTCTCTGTGATGATGATGATGATGATGATGATGATGATGAT 746
Qy      681 GCTGAAGATTCATTCGAAAGAGAGATGCGGTGTTCAACCTTATGCTGATGAGT 740
Db      747 ATCTTACCATGGGAGACAGCAGCAGCAGCCGACCTTAAACCTTCTCTGGAC 806
Qy      741 CAGCGTGGGACAGCATCTCAGAGGTCTCTTGGCTTCAACCTTATGCTGATGAGT 800
Db      807 CTGCGTCTTACTTCTGTGGCAATATCCAGTCCCTGACAAAGAGGCAAGCTGAGCAT 866
Qy      801 GACTTATGTCAGATGCGGTGTGTGGCAGAGCGGTGCTGTGGTATCTCTGTGATCT 860
Db      867 GATGCAATCCAGATATGCAAGCTCGCAGAGAGGAGTGGCGGTATCCGCTGTAGAT 926
Qy      861 GATCCCTTCTCGGTGGCTTGCATGATGATGATGATGATGATGATGATGATGATGATGAT 920
Db      927 GATGCTATGCTTAAAGTGGTCCCTCATCATGCTGCTGTGGCGCATCTCCACCT 986
Qy      921 GGGAGCCAAATGCTGCGGGGTGTGTAACCGAGTCTGGGATTTCCCAAGCTTCAT 980
Db      987 GGGTTTGTATCTGACCCATCTCTGAGTCCCGGCTGCAATCCAGGACATGTGG 1046
Qy      981 CATGGGCTTAACTTCACTTGTGAGTCTGTGATGATGATGATGATGATGATGATGATGAT 1035
Db      1047 CATTAAGATCTGATGAGCATTCCTGGCATCATAGGCGGATGTGGGTGCTGTG 1101

```

RESULT 14
AAH25763
ID AAH25763 standard; cDNA; 1949 BP.

XX AAH25763;

DT 14-AUG-2001 (first entry)

DE Oesophagus cancer associated DRC2 coding sequence.

XX DRC2; oesophagus cancer; ss.

XX Unidentified.

XX Key Location/Qualifiers

FT CDS 41..1480 /tag= a

FT /product= "DRC2"

XX CN1283694-A.

```

PD      14-FEB-2001.
XX      10-AUG-1999; 99CN-00117523.
PF      10-AUG-1999; 99CN-00117523.
PR      10-AUG-1999; 99CN-00117523.
XX      (ONCO-) ONCOLOGY INST TUMOR HOSPITAL CHINESE ACA.
PA      Wang M, Xu Z, Xu X;
XX      WPI; 2001-291757/31.
DR      P-PSDB; AAB97000.
XX      Gene associated with esophagus cancer.
PT      Claim 4; Fig 1; 33pp; Chinese.
XX      The present invention provides the protein and coding sequences of DRC2,
CC      which is associated with esophagus cancer. Also provided is a method for
CC      diagnosing diseases associated with an abnormal version of the nucleic
CC      acid and encoded protein, involving detecting any nucleic acid sequence
CC      mutations, methylation and variation at the RNA and polypeptide level.
CC      The sequences can be used in the treatment of cancer.
SQ      Sequence 1949 BP; 382 A; 593 C; 530 G; 444 T; 0 U; 0 Other;

```

Query Match 13.2%; Score 165.4; DB 5; Length 1949;
Best Local Similarity 49.1%; Pred. No. 9,1e-37;
Matches 499; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

```

Qy      141 CTATCAAGTTGGCCCAAGATCTGACCGTATGCGCGCAATGGGCTTCCCTCAGCTC 200
Db      217 CTACCAAGCTTCCAGAGCGTGCATGATGCTCTCGGAGCTTGGCTTCTCATGAC 276
Qy      201 GAGTTTCCGAGACACACAGCTGAGCAGTGTGACCTTCAACCTTTCATGCTGGCGCTTG 260
Db      277 TTTCTGACGCGCTACGCGCTTACGCGCGGTTCACCTTCTGTTGGAGCCTTGG 336
Qy      261 TGTGAGTGGGCAATCCGCTGAGACCGCTTCTGAGCCAGTTCCTTGTGGAAGTGT 320
Db      337 CATCAGTGGGCGCTGCTCATGACAGGCTGTTCCATCTTCAAGACCGCTTACATCT 396
Qy      321 CATCACTGTTCAGTATTCGCTGGCCACCATGATGCTTTGTGGGTGATCTCAGT 380
Db      397 CGTGGCGGTGAGAACTCATCAAGCTGATCTTGTGGGTGCTTGTGGCGCTT 456
Qy      381 GATGCTGTCTTGGGAGAGTCACTTGGCCAGTGTGTGTGATGATGATGATGATGATGAT 440
Db      457 TGGGCAAGTCTGGGTAAAGTCAAGCCCATTCAGCTCTCATCAATGACTTCTTCAAGT 516
Qy      441 GACAGCTTTAGGCAACCTGAGATGATCATCAGTATATCTTCAAGCAGACTACCAT 500
Db      517 GACCTCTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 576
Qy      501 GAAATGATGACATCTTACGCTTGTGCAAGCTTATTTTGGCTGTGTGGCTGTGGCT 560
Db      577 AGGCTCATGACCATTCACACATTTGGCGCTTACTTTGGGCTGACAGTACCGGATCT 636
Qy      561 GCCAAGCTCTTACCCGAGGGAACGAGAGATTAAGATCAACAGCAGACGATACCACTTT 620
Db      637 CTACCAAGCAACTTAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 696
Qy      621 GTCTGCATGCTGGGCGCTCTTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 680
Db      697 CTTCGCAATGATGGACCTCTTCTGTGATGATGATGATGATGATGATGATGATGATGAT 756
Qy      681 GCTGAAGATTCATTCGAAAGAGAGATGCGGTGTTCAACCTTATGCTGTGACAGT 740
Db      757 ATCTTACCATGGGAGACGAGCAGAGCCGACATCAACCTTATGCTGTGGCAGC 816
Qy      741 CAGCGTGGTACAGCATCTCAGGATCATCTTGTGCTGACCCGCAAGGAGATGAGCA 800
Db      817 CTGCGTCTTACTCTGGGTGCAATATCCAGTGGCTTGCACAAAGAGGAGAGTGAAT 876

```

QY 801 GACTTATGTGCACAGTCCGGTGTGGAGAGCGCTGTGGTACCTGTGTCACT 860
Db 877 GGTCGACATCCAGAAATGSCAAGCTCGCAGAGGGGGTGCCTGGGTACCGCTGTGAGAT 936
QY 861 GATCCCTCTCCCGGCTTCCCATGTGCTGGGTCTTGTGGCTGGGTGATCTTCCGTGG 920
Db 937 GATGCTCATGCTTACCGTCCCTCATCATCGGCTTGTGTGCGCATCATCTCCACCT 996
QY 921 GGGAGCCAACTACCTGCGGGGTGTGTAAACCGAGTCTGGGAGATTCCCAACAGCTTCAT 980
Db 997 GGGTTTGTATACCTGACCCCATTCCTGGAGTCCCGGCTGCACATCCAGACATGTGG 1056
QY 981 CATGGGCTACAACTTCACTGCTGTGGGTCTGCTTGGAGATCATCTTCAATTGTG 1035
Db 1057 CATTAACAATCTGCATGTGCATCTCTGTCATCATAGGCGGAGATGTGGGTGTGTG 1111

RESULT 15
ABK49215
ID ABK49215 standard; cDNA; 1952 BP.
XX
AC ABK49215;
XX
DT 15-JUL-2002 (first entry)
XX
DE cDNA encoding human Rh type C gene (RHCG) protein.
XX
KW RHCG: human; non-erythroid Rh type C glycoprotein; chromosome 15q25;
XX gene; 88.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 24..1440
FT /tag= a
FT /product= "RHCG protein"
FT polyA_signal 1885..1890
FT /*tag= b
XX
PN W0200220719-A2.
XX
PD 14-MAR-2002.
XX
PP 05-SEP-2001; 2001MO-US027503.
XX
PR 07-SEP-2000; 2000US-0230660P.
XX
PA (NYBL-) NEW YORK BLOOD CENT INC.
XX
PI Huang C, Liu Z;
XX
XX WPI; 2002-351774/38.
XX P-PSDB; AAU78997.
XX
PT Nucleic acid sequences encoding novel mammalian nonerythroid Rh type C
PT and glycoproteins which have a characteristic twelve transmembrane domain
PT structure.
XX
XX Claim 2; Fig 1, 53pp; English.
XX
XX This invention relates to the nucleic acid and protein sequences of novel
XX human and mouse non-erythroid Rh type C glycoprotein (RHCG). The RHCG
XX protein and the mouse homologue (rhcg) have a characteristic 12
XX transmembrane domain structure and are expressed in kidneys and testis.
XX The invention also comprises a method for antibody that specifically
XX binds an epitope of the glycoprotein and a method for detecting the
XX protein using this antibody. The antibodies of the invention may be used
XX in Western blot, enzyme linked immunosorbent assays (ELISA) or
XX immunohistochemical assays to identify the non- erythroid tissues,
XX particularly kidney and testis, that express the RHCG or Rhcg
XX glycoproteins. The methods are used for detecting an Rhcg or RhCG
XX glycoprotein in a sample. The present sequence represents the cDNA

CC encoding the human RhCG Rh type C glycoprotein (RhCG) protein sequence of
CC the invention. The gene encoding this protein is located on human
CC chromosome 15q25
XX

SO Sequence 1952 BP; 398 A; 586 C; 524 G; 444 T; 0 U; 0 Other;

Query Match 13.2%; Score 165.4; DB 6; Length 1952;

Best Local Similarity 49.1%; Pred. No. 9.1e-37;
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGGCCAAAGATATGACCGTATGAGGCGCATTTGGGTCTTCTCACTC 200
Db 201 CTACCCAAAGCTTCCAGACGCTGACCGTGAAGTGTCTTGAGGCTTCCCTTCATGAC 260
QY 201 GAGTTTCCGAGACACAGCTGAGAGAGTGGGCTTCAACCTTTCATAGTGGGCTTGG 260
Db 261 TTTCCTGACGCGTACGAGCTTACGCGCTTCAAGCGCTTCAACTTCTGTGGCAAGCTTGG 320
QY 261 TGTGAGTGGGCAATCCTGCTGAGACGGCTTCTGAGCAGATTCCCTTGGGAAAGTGT 320
Db 321 CATCCAGTGGGCGCTGCTCATGACAGGCGTGTCCACTTTCAGAGCCGTAATCGT 380
QY 321 CATCAACATGTTAGTATTTGGCTGGCCACATGAGTCTTGTGGGTGATTCAGT 380
Db 381 CGTGGGCGTGGAGAACCTCATCAACCTGACTTCTGTGGCCCTGTGTGGCGCTT 440
QY 381 GATGCTGTCTTGGGGAGGATCAACTTGGGCGAGTTGGTGGTATGCTGGTGGAGT 440
Db 441 TGGGGCAGTTCTGGTAAAGTCAAGCCCATTCAGCTGTATATAGCTTCTTCCAGT 500
QY 441 GACAGCTTAAAGGCAACCTGAGATGTGATCACTAAATATCTTCAACACAGACTACAT 500
Db 501 GACCTCTTGGCTGTGAATGATTCATCTCTTAACCTGTAAGGTGAAGATGACAG 560
QY 501 GAACATGATGACATATCTACGTGTTCGACGCTTATTTGGGCTGTGTGGCTGTGGCT 560
Db 561 AGGCTCATGACCATCACACATTTTGGCGCTTCTTTGGGCTCACAGTCCCGATCTT 620
QY 561 GCCAAAGCCTCTACCCGAGGGAAGGAGATTAAGATCAAGACAGCAATCCAGTTT 620
Db 621 CTACCGACGCAACTTGAAGACAGACAGAGAGACAAATTTGTGTACAGTGGACCT 680
QY 621 GTCTGCATGCTGGGCGCCCTCTTCTGTGATGTTCTGTGCAAGTTTCAACTCTGCT 680
Db 681 CTTTGCATGATTTGGACCCCTCTTCTGTGATGATCTGGCCAGCTTCAACTACGCA 740
QY 681 GCTGAGAGTCCAAATGAAAGAAATGCGTGTTCACACTTATCTGTAGCACT 740
Db 741 ATCTTACATGGGAGACAGACGACGAGCGCCATCAACACTTACTGTCTTGGGAGC 800
QY 741 CAGCGTGGAGACGCACTCTCAGGTCATCTTGGCTCACCCCAAGGGAAGATCAGCA 800
Db 801 CTGCGTGTCTTACTCTGCTGACAAATATTCAGTCCCTTGCACAAAGAGGCAAGCTGACAT 860
QY 801 GACTTATGTGCACAGTCCGGTGTGGAGAGCGCTGTGGTACCTGTGTCACT 860
Db 861 GGTCGACATCCAGAAATGSCAAGCTCGCAGAGGGGGTGCCTGGGTACCGCTGTGAGAT 920
QY 861 GATCCCTCTCCCGGCTTCCCATGTGCTGGGTCTTGTGGCTGGGTGATCTTCCGTGG 920
Db 921 GATGCTCATGCTTACCGTCCCTCATCATCGGCTTGTGTGCGCATCATCTCCACCT 980
QY 921 GGGAGCCAACTACCTGCGGGGTGTGTAAACCGAGTCTGGGAGATTCCCAACAGCTTCAT 980
Db 981 GGGTTTGTATACCTGACCCCATTCCTGGAGTCCCGGCTGCACATCCAGACATGTGG 1040
QY 981 CATGGGCTACAACTTCACTGCTGTGGGTCTGCTTGGAGATCATCTTCAATTGTG 1035
Db 1041 CATTAACAATCTGCATGTGCATCTCTGTCATCATAGGCGGAGATGTGGGTGTGTG 1095

Search completed: March 23, 2005, 01:08:38
Job time : 709 secs

QY 181 GAGCTTGGGCTTCTCACTGAGATTTCGGAGACAGAGCTGGAGAGAGTGGCTTCAAC 240
DB 181 GAGCTTGGGCTTCTCACTGAGATTTCGGAGACAGAGCTGGAGAGAGTGGCTTCAAC 240
QY 241 CTTTATATGCTGGCGCTTGGTGTGCAAGTGGCAATCTTGTGAGCGGCTTCTGAGCCAG 300
DB 241 CTTTATATGCTGGCGCTTGGTGTGCAAGTGGCAATCTTGTGAGCGGCTTCTGAGCCAG 300
QY 301 TTCCCTTGGGAAAGGTGTGATCACTAGCTTCAAGTATTCGGCTGGCCACCATGAGTGTCT 360
DB 301 TTCCCTTGGGAAAGGTGTGATCACTAGCTTCAAGTATTCGGCTGGCCACCATGAGTGTCT 360
QY 361 TTGTGCTGCTATCTCACTGATGATGCTGCTTGGGAAAGTTCATCTTGGCGCATGTTGGTG 420
DB 361 TTGTGCTGCTATCTCACTGATGATGCTGCTTGGGAAAGTTCATCTTGGCGCATGTTGGTG 420
QY 421 GTGATGCTGCTGATGCTGATGATGCTGCTTGGGAAAGTTCATCTTGGCGCATGTTGGTG 480
DB 421 GTGATGCTGCTGATGCTGATGATGCTGCTTGGGAAAGTTCATCTTGGCGCATGTTGGTG 480
QY 481 TTCAACACAGACTACCAATGACATGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 481 TTCAACACAGACTACCAATGACATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 541 CTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 541 CTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 601 ACAGACAGATACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB 601 ACAGACAGATACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 CCAAGTTTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 CCAAGTTTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 ACTACTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 721 ACTACTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 781 CCCCAGGAGGAGATCAGCAAGACTTATGATGATGATGATGATGATGATGATGATGATGATG 840
DB 781 CCCCAGGAGGAGATCAGCAAGACTTATGATGATGATGATGATGATGATGATGATGATGATG 840
QY 841 GTGGGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 841 GTGGGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 901 GCTGGGCTGATCTCCGTGCGGAGAGCAAGTACCTGCGGGGGTGTGTAACGAGTGTG 960
DB 901 GCTGGGCTGATCTCCGTGCGGAGAGCAAGTACCTGCGGGGGTGTGTAACGAGTGTG 960
QY 961 GGGATTTCCCAAGCTTCATCATGAGGCTCAACTTGTGCTGAGTGTGCTGAGAGAG 1020
DB 961 GGGATTTCCCAAGCTTCATCATGAGGCTCAACTTGTGCTGAGTGTGCTGAGAGAG 1020
QY 1021 ATCATCTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1021 ATCATCTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 TTCCAGGCTCTCTCTGAGATTGGGAACTCAAGTGGCATGTATAGCTTCACTGCT 1140
DB 1081 TTCCAGGCTCTCTCTGAGATTGGGAACTCAAGTGGCATGTATAGCTTCACTGCT 1140
QY 1141 GGTCTCTCAAGAGTTTCTCTTAATCTTAAATATGGAAGCACTTCATGAGGCTTAA 1200
DB 1141 GGTCTCTCAAGAGTTTCTCTTAATCTTAAATATGGAAGCACTTCATGAGGCTTAA 1200
QY 1201 TATTTTGAACCAAGTTTCTGAAAGTTTCTGATTTGGCTGTTGATTTTAA 1254
DB 1201 TATTTTGAACCAAGTTTCTGAAAGTTTCTGATTTGGCTGTTGATTTTAA 1254

RESULT 2
US-09-918-995-30694
; Sequence 30694, Application US/0918995
; Publication No. US20030073622A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30694
; LENGTH: 487
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(487)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30694

Query Match 29.2%; Score 365.8; DB 10; Length 487;
Best Local Similarity 99.5%; Pred. No. 2,4e-106;
Matches 367; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 706 AATGCCGATGTTCAACACCTACTATGCTGTAGCAGTGGGAGAGGACCATCTCAGGG 765
DB 56 AATCCCGTGTTCACACCTACTATGCTGTAGCAGTGGGAGAGGACCATCTCAGGG 115
QY 766 TATCTTGGCTCACCCCAAGGAGATCAGCAAGCTTATGTCACAGTGGCTGTG 825
DB 116 TATCTTGGCTCACCCCAAGGAGATCAGCAAGCTTATGTCACAGTGGCTGTG 175
QY 826 GCAGAGGCGTGGCGTGGGATACCTCGTGTGACCTGATCCCTTCCGTGGCTTGCATG 885
DB 176 GCAGAGGCGTGGCGTGGGATACCTCGTGTGACCTGATCCCTTCCGTGGCTTGCATG 235
QY 886 GTGCTGGGTCTTGTGCTGGCTGATCTCCGTGCGGAGAGCAAGTACCTGCGGGGTG 945
DB 236 GTGCTGGGTCTTGTGCTGGCTGATCTCCGTGCGGAGAGCAAGTACCTGCGGGGTG 295
QY 946 TGTAAACGAGTCTGGGGATTTCCCAACAGCTTCATCATGAGGCTTCAACTTCACTTGTG 1005
DB 296 TGTAAACGAGTCTGGGGATTTCCCAACAGCTTCATCATGAGGCTTCAACTTCACTTGTG 355
QY 1006 GGTCTGCTTGGAGATCATCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1065
DB 356 GGTCTGCTTGGAGATCATCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 415
QY 1066 AATGGCATG 1074
DB 416 AATGGCATG 424

RESULT 3
US-10-302-172-362
; Sequence 362, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: drmanac, Radjoe T
; TITLE OF INVENTION: Polypeptides
; TITLE OR INVENTION: No. US20040053250A1e1 Arginine-rich Protein-like Nucleic Acids and
; FILE REFERENCE: 803, 1CNCB
; CURRENT APPLICATION NUMBER: US/10/302,172
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20

PRIOR APPLICATION NUMBER: PCT US02/05095
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 09/739,451
PRIOR FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 950
SOFTWARE: PC_FIL_genes Version 2.0
SEQ ID NO 362
LENGTH: 1792
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (32)..(1354)
US-10-302-172-362

Query Match 14.0%; Score 175.4; DB 17; Length 1792;
Best Local Similarity 49.4%; Pred. No. 6.6e-45;
Matches 455; Conservative 0; Mismatches 466; Indels 0; Gaps 0;

141 CTATCAAGTTGGCCAGATCTGACCGTGATGCGGCCATTGGCGCTTGGCTCTACCTC 200
211 CTACCCAGCTTCAGAGACGTCATGCTGCTTGTGGCGCTTACCTTCTCATGCT 270
201 GAGTTTCGAGACACAGCTGAGACGTGGCCCTTCAACCTTTCACTGCGCGCTTGG 260
271 CTTCCTGACAGCTTACGCTTTCAGACAGCTGGCGCTTCACTTCTCTGCGCGCTTGG 330
261 TGGGAGTGGGCAATCCTGCTGAGCGGCTTCTGAGCAGTTCCTCTGGGAGAGTGT 320
331 CCGTAGTGTTCACACTGTGCTCCGCGCTTCTCACTCTCTCCAGCGTGGCCATCCA 390
321 CATCACACTGTAGTATTCGGCTGCGCACCATGAGTGTTCGCTGCTGATCTGAT 380
391 TGTGGGTGGAGAGCATGATCATGCTGACTTTTGTGCGGGGCCGCTGCTCATCTCT 450
381 GGAATGCTGCTTGGGAGAGTCAACTTGGCGCACTGTGGTGTGCTGCTGAGGT 440
451 TGGTGGCGCTGCTGGCAAGACCGGGCTTACCACTGCTGCTGAGCCCTGCTGAGGT 510
441 GACAGCTTGGGCACTGAGAGTGCATCATGATATCTTCAACACAGACTACCAT 500
511 GGTGCTGTTGGCATCAATGATGTTGCTCTTCACTCTCTGGGGGTGAGAGTGGCG 570
501 GAACATGATGACACTTACGCTGTCGAGCCTATTTTGGCGTGTCTGTGGCTGTGCT 560
571 AGGCTCATATACATTCACACACTTGTGGTCTACTTGGGGTCTCTTTTGGCGGTCT 630
561 GCCAAGCCTCTACCCGAGGGAACGAGATTAAGATCAACAGCAACGATACCACTT 620
631 GTACAGGCCCGAGCTGAGAAAGACACCGCAGGGCTCCGTTACATTCAGACT 690
621 GTCTGCATGCTGGGGCCCTCTTCTGTGATTTCTGGCGCAAGTTTCAACTTGTCT 680
691 CTTCGCATATTTGGGACATCTTCTGTGATTTCTTGGCTTGTGCTCAATGCTGACT 750
681 GCTGAGAAGTCAATTCAGAGAAAGATGCCGTGTCAACCTACTATGCTGAGAGT 740
751 CACAGGCTGGGGCTGGGACGATCGAAGCGCCCTCAACACATACATCTCTGGCTGC 810
741 CAGGCTGTGACACGCCATCTCAGGGTATCTTGGCTTACCCCGAGGGAAGATCAAC 800
811 CAGACACCTTGGGACCTTTGGCTTGTGACGCCCTTGTAGGGGAAGATGGAGGCTTGA 870
801 GACTTATGTGACAGTGGGTGTGGGAGAGGCGTGTGGTGTACCTGTGTCACT 860
871 GGTTCACATCCAAATACAGCGCTGGCTGAGAGGGGTGTGTGGGAGCTTCAATGTA 930
861 GATCCCTTCTCCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
931 GATCTTACACCTTTTGGGGCTCTGGGAGCTGCTTCTTGGCTGGGAGTGTCTCAAGCT 990
921 GGGAGCAAGTACCTGCGGGGTGTGTAAACGAGTCTGGGAGTTCCCAACAGCTCAT 980

DB 991 GGGGACAGTCTTCAACGCCCATCTTGATCAAAATTCAAAAGTCCAGACATGTGG 1050
QY 981 CATGGCTACAACTTCACTTGTGCTGGGTCTGTGGAGAGATCATCTACATGTGCTGT 1040
DB 1051 AGTCCACAACTTCCATGAGATGCGGGGCTCTGGGGGCTCTGGGGGCTCTGTGGC 1110
QY 1041 GGTGCTGATACCTGCGAGC 1061
DB 1111 TGGACTTGCACCCCATGAGAC 1131

RESULT 4

US-10-074-978A-41/C
Sequence 41, Application US/10074978A
Publication No. US20040010119A1

GENERAL INFORMATION:

APPLICANT: Leite, Mario
APPLICANT: Spytek, Kimberly A
APPLICANT: Guo, Xiaojia (Sasha)
APPLICANT: Fernandes, Bima
APPLICANT: Li, Li
APPLICANT: Kekuda, Ramesh
APPLICANT: Liu, Xiaohong
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc
APPLICANT: Patuturajan, Meera
APPLICANT: Blalock, Angela
APPLICANT: Ballinger, Robert
APPLICANT: Vermet, Corine
APPLICANT: Tchernev, Velizar T
APPLICANT: Malyankar, Urfel M
APPLICANT: Gusev, Vladimir
APPLICANT: Rastelli, Luca
APPLICANT: Mezes, Peter S
APPLICANT: Ellerman, Karen
APPLICANT: Heyes, Melvin P
APPLICANT: Heitman, John
APPLICANT: Pena, Carol E A
APPLICANT: Shinkets, Richard A
APPLICANT: Taupier Jr, Raymond J
APPLICANT: Moore, No. US20040010119A111e
APPLICANT: Shenoy, Suresh
APPLICANT: Edinger, Shlomil
APPLICANT: Gunther, Erik
APPLICANT: Stone, Dave
APPLICANT: Miller, Isabelle
APPLICANT: Beyman, John
APPLICANT: Smithson, Glenda
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-269
CURRENT APPLICATION NUMBER: US/10/074,978A
CURRENT FILING DATE: 2003-01-07
PRIOR APPLICATION NUMBER: 60/268,221
PRIOR FILING DATE: 2001-02-12
PRIOR APPLICATION NUMBER: 60/335,109
PRIOR FILING DATE: 2001-10-31
PRIOR APPLICATION NUMBER: 60/312,284
PRIOR FILING DATE: 2001-08-14
PRIOR APPLICATION NUMBER: 60/268,496
PRIOR FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/276,703
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/330,293
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/322,127
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: 60/280,899
PRIOR FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: 60/310,797
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: 60/268,646
PRIOR FILING DATE: 2001-02-14
Remaining Prior Application data removed - See File Wrapper or PALM.

Db 897 GATGCTCATGCTTACCGTGCCTTCATCGGCTTGTCTGCGGACATCATCTCCACCT 956
Qy 921 GGGAGCCAGAACCTGCGGGGGTGTGAAACGAGTGTGGGGATTCGCCACAGCTCAT 980
Db 957 GGGTTTGTATACCTGACCCCATTCCTGAGAGTCCCGGCTGCACATCCAGACATATGCG 1016
Qy 981 CATGGGCTACAACTTCAGCTTGTGGTGTCTGCTTGAGAGATCATCTACATTTGTG 1035
Db 1017 CATTAACATCTGATGACATCTTCCTGACATCATAGGGGACATCGTGGGTGCTGTG 1071

RESULT 6

us-09-949-145-1
; Sequence 1, Application US/09949145
; Patent No. US20020055622A1
; GENERAL INFORMATION:
; APPLICANT: New York Blood Center
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyc
; FILE REFERENCE: Docket 454-31
; CURRENT APPLICATION NUMBER: US/09/949,145
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/230660
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1952
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AF193809
; DATABASE ENTRY DATE: 1999-12-22
; RELEVANT RESIDUES: (1)..(1952)
US-09-949-145-1

Query Match 13.2%; Score 165.4; DB 9; Length 1952;
Best Local Similarity 49.1%; Pred. No. 1.1e-41;
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

Qy 141 CTATCAAGTTGGCCAAATCTGACCGTGATGGCGCCATTGGGCTTCTTCACCTC 200
Db 201 CTACCCAGCTTCCAGACGCGTGAATGCTTGTGGGCTTCCGCTTCATGAC 260
Qy 201 GAGTTCCGGAGACACAGCTGAGAGTGGGCTTCAACCTCTTCAAGTGGGCTTGG 260
Db 261 TTTCCTCAGCGCTACGAGCTTCAAGCGCGGCGGCTTCACTTCTGTGGACGCTTGG 320
Qy 261 TGTGCAATGGGCAATCTGCTGACGCGCTTCTGAGCCAGTCCCTTCTGGAGAGTGT 320
Db 321 CATCCAGTGGGCGCTGCTCATGAGGCGCTGTTCCATCTTCAAGACGCTACATCGT 380
Qy 321 CATCACTGTTCAATTCGCTGCGCCACCATGATGCTTGTGGTGTGATCTGAGT 380
Db 381 CGTGGGCGTGAAGAACCTCAACGCTGACTTCTGGGCGCTCTGTCTGCGGCGCTT 440
Qy 381 GATGCTGTCTGGGGAAGGCACTTGGCGCACTTGTGTGATGTGCTGTGAGAGT 440
Db 441 TGGGGCGTTCTGGTAAAGTCAAGCCCATTCACCTCTCAATGACTTCTTCCAGT 500
Qy 441 GACAGCTTTAGGCAACCTGAGATGTCATCAGTAATATCTTCAACACAGACTACCAT 500
Db 501 GACCCCTTGTGGCTGTGATGATGTTCTATCTTCTTAACTGTGTAAGTGAAGATGACG 560
Qy 501 GAACATATGACATCTTACGCTGTTCGACGCTTATTTTGGGCTGTCTGTGGCTGTGCT 560
Db 561 AGGCTTCATGACCTTCAACATTTGGGCGCTTACTTGGGCTCAAGTGAACCCGATCT 620
Qy 561 GCCAAAGCTTCAACGAGGGAAGGATTAAGATCAGACAGCAAGATACCAAGTTT 620
Db 621 CTACCAAGCAACTTAAGAGAGAGCAAGAGACAGATTTCTGTACAGTGAAGCT 680
Qy 621 GTCTGCATGTGCGGCGCTCTTCTGTGTGATGTTCGCAAGTTTCAACTCTGCTCT 680

Db 681 CTTGCGATATTTGGACCCCTTCTGTGATGTACTGGGCCAGCTTCAACTCAGCCAT 740
Qy 681 GCTGAGAACTCAATGAAAGAAAGAAATGCGGTGTTCAACACTTATGCTGAGAGT 740
Db 741 ATCTTACCATGGGAGACGCCAGCACCGAGCGCCATTAACACTTACTGCTCTTGGCAGC 800
Qy 741 CAGCGTGTGACAGCCCATCTCAAGGTCATCTTGGCTCACCCCCAGAGATCAGCA 800
Db 801 CTGGGTCTTACCTCGGCTGCAATATCAGATGCCCTGCACAAAGAGGGCAAGTGCAT 860
Qy 801 GACTTATGTGACAGTCCGCTGTGGCAGAGGCGCTGTGGGTACCTGTCTACCT 860
Db 861 GGTGCATCATCAGAAATGCCAGCTCGCAGAGGGGTGCGCTGGGTACCTGTCTGAGT 920
Qy 861 GATCCCTTTCGCGGCTTGCATGCGTGGGCTGTGGGCTGTGGGCTGTGATCTCGTGG 920
Db 921 GATGCTATGCTTACCGTGCCTTCATCATGAGTCTTGTGCGGCAATCTTCACCT 980
Qy 921 GGGAGCCAAATACCTGCGGGGTTGTGTAACGAGTCTGGGAGTTCCACACAGCTTCAT 980
Db 981 GGGTTTGTATACCTGACCCCATTCCTGAGAGTCCCGCTGCACATCCAGACATATGCG 1040
Qy 981 CATGGGCTACAACTTCAGCTTGTGCTGTGCTGTGAGAGATCATCTTACATTTGTG 1035
Db 1041 CATTAACATCTGATGACATCTTCCTGACATCATAGCGGACATCGTGGGTGCTGTG 1095

RESULT 7

US-10-172-118-1635
; Sequence 1635, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1635
; LENGTH: 1952
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_016321
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1635

Query Match 13.2%; Score 165.4; DB 17; Length 1952;
Best Local Similarity 49.1%; Pred. No. 1.1e-41;
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

Qy 141 CTATCAAGTTGGCCAAATCTGACCGTGATGGCGCCATTGGGCTTCTTCACCTC 200
Db 201 CTACCCAGCTTCCAGACGCTGACGATGATGTCTTGTGGGCTTCCGCTTCTTCATGAC 260
Qy 201 GAGTTCCGGAGACACAGCTGAGAGTGGGCTTCAACCTTCTTCAATGCTGGGCTTGG 260
Db 261 TTTCCTCAGCGCTACGAGCTTCAAGCGCGGCTTCACTTCTGTGGCAGGCTTGG 320
Qy 261 TGTGCAATGGGCAATCTGCTGACGCGCTTCTGAGCAATTCCTTCTGTGGAAGTGT 320
Db 321 CATCAATGGGCGCTGCTCATGAGGCGGCTGTTCATCTTCTTCAAGACGCTACATCGT 380

```

Qy 321 CATCACTGTTAGTAATTCGGCTGGCCACATAGTGGCTTTGCGGTGCTGATCTCACT 380
Db 381 CGTGGGGGTGGAGAACCTTCATCAACGCTGACTTCTGCGTGGCTCTGTCTGGGTGGCTT 440
Qy 381 GGATGCTGTCTTGGGGAGGTCAACTTGGCGCATTTGGTGGTGTGATGTCTGGTGGAGT 440
Db 441 TGGGGCAGTTCTGGTGAAGTCAAGCCCCATTACGCTGCTCATATGACTTTCTTCCAAAGT 500
Qy 441 GACAGCTTTAGGCACTGAGATGATCAATGATATATCTTCAACAGACTTACCAACT 500
Db 501 GACCTCTTCTGCTGATGATGATTCATTCCTTAACTGCTTAAAGTGAAGATGACAG 560
Qy 501 GAACATGATGACATCTAGCTGTGGACACCTTATTTGGCTGTGTGGCTGTGGCT 560
Db 561 AGGCTTCATGACCATTCACACATTTGGCCCTTACTTTGGCTTCACAGTGAACCCGATCTT 620
Qy 561 GCCAAGCTCTTACCCGAGGAACGAGGATTAAGTATCAAGACAGAACTACCAATTT 620
Db 621 CTACCGAGGCAACTTGAAGCAGACAGAGAGACAGAAATTTGTGTACCACTCGACCT 680
Qy 621 GTCTGCTATGCTGGGGGCTCTTCTTGTGATGATTTGAGCCAGTTTCAACTGCTCT 680
Db 681 CTTTGCATGATTTGGCACCTCTCTGTGATGTACTGGCCAGCTTCAACTGACCAT 740
Qy 681 GCTGAGAGTCCATTCAGAAAGGAAGATGCGGTTCACACCTTACTATGCTGTAGCACT 740
Db 741 ATCTCAACATGGGGGACAGCAGCAGCCAGCCGATCAACAATCTGCTCTTGGCAGC 800
Qy 741 CAGCGTGTGACAGCCATCTCAGGGTCACTCTTGGCTCACCCCGAAGGAAGATCAGCA 800
Db 801 CTGCTGTCTTACCTCGGTGGCAATTCAGATGCTCTGACAAAGAGGCAAGCTGACAT 860
Qy 801 GACTTATGTGACAGTGGGTGTGTGGCAGAGGGGTGGCTGTGGGTACTGTGTCACT 860
Db 861 GGTGACATCCAGATGACAGCTGCGCAGAGGGGTGGCTGTGGGTACTGTGTCACT 920
Qy 861 GATCCCTTCTCCGCTGTGGCTGATGCTGTGGGTGTGTGGCTGTGGGTACTGTGTCACT 920
Db 921 GATGCTCATGCTTACGCTGCTGCTCATCATGCTGTGTGTGTGTGTGTGTGTGTGTGT 980
Qy 921 GGGAGCCAAAGTACCTGCGGGGTGTGTAAACGAGTGTGGGATTTCCCAAGCTTCAT 980
Db 981 GGGTTTGTATACCTGACCCCATTCCTGAGTCCCGGCTGCACATCCAGAGACATGTGG 1040
Qy 981 CATGGGCTACAACTTCACTTGTGCTGGGTGCTTGGAGATCATCTACATTTGTG 1035
Db 1041 CATTAACAATCTGATGATGCAATCTGATCATATAGGGGCACTGTGGGTGTGTG 1095

```

RESULT 8
US-10-342-887-1635
Sequence 1635, Application US/10342887
Publication No. US20040058340A1
GENERAL INFORMATION:

APPLICANT: Dai, Hongyue
APPLICANT: He, Yudong
APPLICANT: Linsley, Peter S.
APPLICANT: Mao, Mao
APPLICANT: Roberts, Christopher J.
APPLICANT: Van 't Veer, Laura Johanna
APPLICANT: Van de Vijver, Marc J.
APPLICANT: Bernards, Rene
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
FILE REFERENCE: 9301-188-999
CURRENT APPLICATION NUMBER: US/10/342,887
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/298,918
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/380,710
PRIOR FILING DATE: 2002-05-14
PRIOR APPLICATION NUMBER: 10/172,118
PRIOR FILING DATE: 2002-06-14
NUMBER OF SEQ ID NOS: 2699

SEQ ID NO 1635
LENGTH: 1952
TYPE: DNA
ORGANISM: Homo sapiens
US-10-342-887-1635

Query Match 13.2%; Score 165.4; DB 17; Length 1952;
Best Local Similarity 49.1%; Pred. No. 1.1e-41;
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

```

Qy 141 CTATCAAGTTGGCAAGATCTGACCGTGAATGGCGGCATTTGGTGGCTTCTTCACTTC 200
Db 201 CTATCAAGTTGGCAAGATCTGACCGTGAATGGCGGCATTTGGTGGCTTCTTCACTTC 260
Qy 201 GAGTTTCCGAGACACAGCTGTGAGAGTGTGCTTCAACTCTTCACTATGCTGGCTGG 260
Db 261 TTTCTGACAGCGCTACGCTTACGCGCTGGGCTTCAACTCTTGTGTGGAGCTTGG 320
Qy 261 TGTGAGTGGGCAATCTGCTGAGACGCTTCTTCAAGCCAGTTTCTTGGGAAGTGT 320
Db 321 CATCACTGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 380
Qy 321 CATCACTGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 380
Db 381 CGTGGCGTGGAGAACCTCATCAAGCTGATCTTGTGCTGCTGTGTGTGTGTGTGTGT 440
Qy 381 GATGCTGTCTTGGGGAAGTCACTTGGCGCATTTGTGTGTGTGTGTGTGTGTGTGTGT 440
Db 441 TGGGGAGTTTGGGTAAAGTACGCCCATTCAGCTGCTCATATGACTTTCTTCAAGT 500
Qy 441 GACAGCTTTAGGCACTGAGATGCTGATGATTAATCTTCAACACAGACTACACAT 500
Db 501 GACCTCTTCTGCTGATGATGATTAATCTTCAACACAGACTACACAT 560
Qy 501 GAACATGATGACATCTAGCTGTGGACACCTTATTTGGGTGTGTGTGTGTGTGTGTGT 560
Db 561 AGCTTCAATGACCATTCACATTTGGGCTTACTTTGGGTGTGTGTGTGTGTGTGTGT 620
Qy 561 GCCAAGCTCTTACCCGAGGAACGAGGATTAAGTATCAAGACAGAACTACCAATTT 620
Db 621 CTACCGAGGCAACTTGAAGCAGACAGAGAGACAGAAATTTGTGTACCACTCGACCT 680
Qy 621 GTCTGCTATGCTGGGGGCTCTTCTTGTGATGATTTGAGCCAGTTTCAACTGCTCT 680
Db 681 CTTTGCATGATTTGGCACCTCTCTGTGATGTACTGGCCAGCTTCAACTGACCAT 740
Qy 681 GGTGACATCCAGATGACAGCTGCGCAGAGGGGTGGCTGTGGGTACTGTGTCACT 740
Db 741 ATCTCAACATGGGGGACAGCAGCAGCCAGCCGATCAACAATCTGCTCTTGGCAGC 800
Qy 741 CAGCGTGTGACAGCCATCTCAGGGTCACTCTTGGCTCACCCCGAAGGAAGATCAGCA 800
Db 801 CTGCTGTCTTACCTCGGTGGCAATTCAGATGCTCTGCTGCTGCTGCTGCTGCTGCT 860
Qy 801 GACTTATGTGACAGTGGGTGTGTGGCAGAGGGGTGGCTGTGGGTACTGTGTCACT 860
Db 861 GGTGACATCCAGATGACAGCTGCGCAGAGGGGTGGCTGTGGGTACTGTGTCACT 920
Qy 861 GATCCCTTCTCCGCTGTGGCTGATGCTGTGGGTGTGTGGCTGTGGGTACTGTGTCACT 920
Db 921 GATGCTCATGCTTACGCTGCTGCTCATCATGCTGTGTGTGTGTGTGTGTGTGTGTGT 980
Qy 921 GGGAGCCAAAGTACCTGCGGGGTGTGTAAACGAGTGTGGGATTTCCCAAGCTTCAT 980
Db 981 GGGTTTGTATACCTGACCCCATTCCTGAGTCCCGGCTGCACATCCAGAGACATGTGG 1040
Qy 981 CATGGGCTACAACTTCACTTGTGCTGGGTGCTTGGAGATCATCTACATTTGTG 1035
Db 1041 CATTAACAATCTGATGATGCAATCTGATCATATAGGGGCACTGTGGGTGTGTG 1095

```

RESULT 9
US-09-949-145-6

```

Sequence 6, Application US/09949145
Patent No. US20020055622A1
GENERAL INFORMATION:
APPLICANT: New York Blood Center
TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyco
FILE REFERENCE: Docket 454-31
CURRENT APPLICATION NUMBER: US/09/949,145
CURRENT FILING DATE: 2001-09-07
PRIORITY APPLICATION NUMBER: 60/230660
PRIORITY FILING DATE: 2000-09-07
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 1497
TYPE: DNA
ORGANISM: Mus musculus
US-09-949-145-6

```

Query Match	12.0%	Score 150.2;	DB 9;	Length 1497;
Best Local Similarity	47.5%	Pred. No. 7.4e-37;		
Mismatch 446; Conservative	0;	Mismatches 493;	Indels 0;	Gaps 0;

QY	97	CACHTAGCGCTTCTCTTAGAGATCAAAAGGCGCTCGTGGCATCTTAAAGTTGGCCA	156
Db	136	CGAAGAACATCTCCAGCGACGTTGAGAACAGATTTTATCTATCGCTACCAGCTTCAG	195
QY	157	GATCGACCGGTGATGAGGGGCCATTGGGCTTGGCTCCACCTGAGTTTCCGAGAC	216
Db	196	GATGTACACGCCATATGTTCTTGTTGGGCTTCGGCTTCTCTATGACCTTCTCAGAGCTAC	255
QY	217	AGCTGAGACAGTGTGGCCCTTCAACCTCTTCAATGCTGGCGCTTGTGTGACAGTGGACATC	276
Db	256	GGCTTCAGCGCTGATAGGCTTCAACTCTGCTGGGACGCTTTGGGATCAGTGGGCACTG	315
QY	277	CTGCTGACGCGCTTCTGAGCCAGTCCCTTCTGGAAAGGTGTATCACTGTTCACT	336
Db	316	CTCATGCAAGGATGGTGTCCATTACTTTGAMAGAGCCACTTGTCTGAGGCGTGAAGAC	375
QY	337	ATTGGGCGGGCCACCATGATGCTTGTGCGGTGATCTGAGGGATGCTGTCTTGGGG	396
Db	376	ATCATCCAACTGATTTCTGTGTGGCATCTTCTGTGTGGCTTGGGGGAGTTCTTAGGC	435
QY	397	AAGGTCAACTTGGCGCAGTGTGTGATGGTGTGCTGTGAGAGTACACGTTTAGGCAC	456
Db	436	AAAGTCAGCCCGATGACGCTGCTCATTTATGACCTTCTTCAAATGATCTCTTCAAGTG	495
QY	457	CTGAGGATGTGATCATGTAATATCTTCAACAAGACTACCAATGAACATGATGACATC	516
Db	496	AATGATGTCATCTCTCTGACCTGTAGAGCCAAAGATGCAAGGGGGCTCTATGACCATC	555
QY	517	TACGTGTCGAGGCTATTTTGGGCTGTCTGTGGCTGTGATGCTGCCAAACCTCTACCC	576
Db	556	CACACATTTGGCGCTTACTTTGGGCTCAAGAGACTGTGAATCTCTACCGAAAAAAGCTTG	615
QY	577	GAGGGAACGAGATTAAGATCAGACCAACGATACCCAGTTGTCTGCCATGCTGGGC	636
Db	616	GATCAGACCAAGCAGAGACAGAGCTCAGTGTACCACTCGACCTTTTGGCATGTTGGC	675
QY	637	GCCCTCTTCTTGTGTGATGTTTGTGGCCAAAGTTTCAACTCTGCTCTGTGTAAGATTCATC	696
Db	676	ACCCCTCTTCTGTGTGATTAATCTGGGCCAGTTTCAATTCAGCAGTTCCCTTCCACGAGAT	735
QY	697	GAAAGGAAGAAATGCCGTGTTCAACACCTATGTGTGAGAGCTGAGGTGTGTACAGCC	756
Db	736	GCCACGACCCGAGCAGCCCTCAATACCTAATCTTCTTGGCAGGAGTGTCTAACACA	795
QY	757	ATCTCAGGGTCACTCTTGGCTACCCCCAAAGGAAAGATCAGACAGACTTATGTGTACAGT	816
Db	796	GTGACAGTATCCAGTATTTGTACACAAGAAGGCAAGTTGATATGTGTCAATCCAGAT	855
QY	817	GGGGTGTGGCAGAGGCGGTGGCTGTGGTATCTCTGTCACTGTATCCCTTCTTCGCTG	876
Db	856	GCCACGCTTTCAGAGTGTGGGTGTGGGCAAGCTGTGGAGATGATCTCTCAACCTTAC	915

QY	877	CTTGCCATGAGTCTGGGCTCTTGGAGCTGGGCTATCTCCGTCGGGGGAGCCAAAGTACTG	936
Db	916	GAGCGCTCATCGAGGGGTTCTTCTGGGGCAATTTTCCACCCTAGAGTTTGCACTACTTA	975
QY	937	CCGGGGGTGTTAACCGAGTCTGGGGATTCCCAAGCTTCATCATGGGCTACAATTC	996
Db	976	ACGCCATTCTCGAGTCCCGGCTTCGATCCAGGACATGTGGCATTCACAACTGCAC	1035
QY	997	AGCTTGTGGGTCTGCTTGGAGAAATCATCTACAATTGTG	1035
Db	1036	GGCATTCCTCGCATCATAGGGGCGCATGTGTGGTCTGTG	1074

RESULT 10
US-09-949-145-2
: Sequence 2. Application US/09949145

```

Patient No. US20020055622A1
General Information:
Applicant: New York Blood Center
Title of Invention: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyco
File Reference: Docket 454-31
Current Application Number: US/09/949,145
Current Filing Date: 2001-09-07
Prior Application Number: US 60/230660
Prior Filing Date: 2000-09-07
Number of SEQ ID NOS: 77
Software: PatentIn version 3.1
SEQ ID NO 2
Length: 2097
Type: DNA
Organism: Mus musculus
Publication Information:
Database Accession Number: AF193810
Database Entry Date: 1999-12-22
Relevant Residues: (1)..(2097)
US-09-949-145-2

```

Query Match	12.0%;	Score 150.2;	DB 9;	Length 2097;
Best Local Similarity	47.5%;	Pred. No. 8.8e-37;		
Matches 446; Conservative	0;	Mismatches 493;	Indels 0;	Gaps 0;

QY	97	CACRTAGACGCTTCTTAGAGATCAAAAGGGGCTCGTGACATCTTATCAAGTTGGCCAA	156
Db	258	CGAAGAACATCTCCAGGACGTTGAGAACAAAGATTTTATCTATCGCTAACCGAGCTTTCAG	317
QY	157	GATCGACCGGTGATGAGCGGCACATGGGCTTTGGGCTTCTCTCACTCGAGTTTCCGAGACAC	216
Db	318	GATGTACACGCCATGCTTCTGTGGGCTTCGGCTTCTCTCAAGACCTTCTCTCAGGCGTAC	377
QY	217	AGCTGAGACAGTGTGGCCTTCAACCTCTTCAATGCTGCGCTTGGTGTGACGTGGCAATC	276
Db	378	GGCCTTCAGCGCGTGAAGGCTTCAACTCTCGCTGGCAGCTTTCGGGATTCAGATGGGCACTG	437
QY	277	CTGCTGAGCGGCTTCGTAGACGACGATGCCCTTCTGAGGAAGGTGATCAGACTGTTCAGT	336
Db	438	CTCATGCAAGGATGGTTCATTATCTTTGAAAGAGGCCAATTGTCCTAGGCTCGAGAAC	497
QY	337	ATTGCGCTGGCCACCATAGTGTCTTGTGCGGTGATCTTCAGTGAATGCTGTCTTTGGG	396
Db	498	ATCATTCGAAGCTAATTCTGTGTGGCACTTCTCTGTGTGGCCTTCGGGGCAGTTCTAAGC	557
QY	397	AAGGTCAACTTGGCGCAGTTGGTGTATGTGTGCTGTGTGAGGTGACAGCTTTAGGCAAC	456
Db	558	AAGGTCAAGCCGATGACAGCTGCTCAATTAATGACCTTCTTCAAGTGAATCTTCTTCAACAGT	617
QY	457	CTGAGAGTGTGATCAGTAATATCTTCAACACAGACTACCAATGAACATGATGCACATC	516
Db	618	AATAGATTCAATCTCTCGAACCCTATATAGGCCAAAGATGCAAGGGGGCTTATGACATC	677
QY	517	TACGTGTTGCGAGCCTATTTTGGGCTGTCTGTGGCTGTGTGCTGCTGCCAACGCTTACCC	576
Db	678	CACACATTTTGGCGCCTTACTTTTGGGCTCAACATGACCTGGATCTCTTACCGAAAAACCTG	737

```
QY 577 GAGGAGAGGAGATTAAGATCAGACAGACATACCAGTTTGTCTGACATGCTGGGC 636
DB 738 GATCAGAGACAGACAGACAGACCTCAGTGTACCACTCGACCTTTTGGCCATGTTGGC 797
QY 637 GGCCTCTCTTGTGATGTTCTGCGCAAGTTTCAACTCTGCTCTGCTAGAGATTCATC 696
DB 798 ACCCTCTTGTGATGATTAAGTGGCCAGTTTCAATTCAGCAAGTTCCTTCACGAGAT 857
QY 697 GAAAGAGATACCGATGTTCAACCTATATGCTAGACAGTGGTGTGACAGCC 756
DB 858 GCCCAGCACCCAGACAGCCCTCAATCTATCTCTTGGCAGGAGTGTCTTACCA 917
QY 757 ATCTAGAGTCACTCTTGGCTTCAACCCCAAGGAAAGATCAGCAAGATTATGACAGT 816
DB 918 GTGACAGTATCCAGTATGTATACAAAGAGGCAAGTTGATATGTCACATCCAGAT 977
QY 817 GCGGTGTTGAGAGAGGCGGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 876
DB 978 GCCAAGCTTGCAGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1037
QY 877 CTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 936
DB 1038 GACCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1097
QY 937 CCGGGGTGTTGAACCAAGTCTGAGGATTTCCCAAGCTTCATGAGGCTACCACTTC 996
DB 1098 ACGCATTCCTGAGTCCCGCTTCGATCCAGACATATGCAATTCACCAACTGCAC 1157
QY 997 AGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1035
DB 1158 GGCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1196
```

```
RESULT 11
US-10-047-730/c
; Sequence 730, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT FILING DATE: 2002-03-25
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 730
; LENGTH: 2415
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-730
```

```
Query Match 11.9%; Score 148.8; DB 17; Length 2415;
Best Local Similarity 65.9%; Pred. No. 2,7e-36;
Matches 216; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
```

```
QY 758 TCTCAGAGTCACTTGTGCTACCCCAAGGAAATCAGCAAGATTATGTCAGAGT 817
DB 1875 TCTCTACCTGCTTCTTACCCCAAGCTATTTCTTGCAAGACTTATGTCAGAGT 1816
QY 818 CCGTGTGAGAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 877
DB 1815 CCGTGTGAGAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1756
QY 878 TTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 937
DB 1755 TTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1696
QY 938 CCGGGGTGTTGAACCAAGTCTGAGGATTTCCCAAGCTTCATGAGGCTACCACTTCA 997
DB 1695 CCGTAAAGAACTGAACACTTCTCTGCTTGTGCTGAGAGGCAAGACGCTGGG 1636
```

```
QY 998 GCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1057
DB 1635 ACCGTATGGGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1576
QY 1058 GAGCCGCAATGAGCATGATGTTGGCTTCA 1085
DB 1575 TTGGCTTCAACGCTGATGAGGATTCGA 1548
```

```
RESULT 12
US-10-027-632-282039
; Sequence 282039, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 282039
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-282039
```

```
Query Match 11.5%; Score 144.6; DB 13; Length 505;
Best Local Similarity 86.9%; Pred. No. 2,7e-35;
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
```

```
QY 758 TCTCAGAGTCACTTGTGCTACCCCAAGGAAATCAGCAAGATTATGTCAGAGT 817
DB 241 TCTCTACCTGCTTCTTACCCCAAGCTATTTCTTGCAAGACTTATGTCAGAGT 300
QY 818 CCGTGTGAGAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 877
DB 301 CCGTGTGAGAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 878 TTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 937
DB 361 TTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 938 CGG 940
DB 421 CGG 423
```

```
RESULT 13
US-10-027-632-282040
; Sequence 282040, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
```


Db	361	TTGCCATGGTGCTGGGCTTTGTGGCTGGGCTGATCTCCGTCGGGGAGCCAACTACTGC	420
Qy	938	CGG	940
Db	421	CGG	423

Search completed: March 23, 2005, 07:24:10
Job time : 4489 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2005, 23:43:09 ; Search time 3942 seconds
(without alignments)
12108.725 Million cell updates/sec

Title: US-09-600-714-41
Perfect score: 1254
Sequence: 1 atgagctcaagaccgcgcg.....attgctgtgtgattctaa 1254

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_hlc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gse1.*
9: gb_gse2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	824.4	65.7	964	5	BX406635
2	633.8	50.5	685	5	BU664461
3	629.8	50.2	676	5	BX106517
4	617.2	49.2	685	5	BU655978
5	602.2	48.0	1457	3	AK079335
6	597.4	47.6	706	5	BU661907
7	585.4	46.7	878	5	BX406634
8	560.4	44.7	631	5	BU662207
9	559	44.6	632	5	BU664450
10	552	44.0	613	5	BU664659
11	548.8	43.8	888	7	CK453280
12	542.4	43.3	607	5	BU657533
13	498.6	39.8	975	3	AK089642
14	429.2	34.2	458	6	CD627633
15	427.2	34.1	612	5	BU663660
16	418.6	33.4	862	7	CK775698
17	410	32.7	461	6	CD627634
18	405.6	32.3	587	5	BU661598
19	399.2	31.8	480	5	BU659596
20	398.6	31.8	683	4	BM719724
21	392.4	31.3	443	4	AL699045
22	377.8	30.1	873	7	CN166957
23	375.6	30.0	870	7	CK452223
24	366	29.2	804	7	CO807980

25	359	28.6	429	5	BU657883	BU657883 c129d10.z
26	349.2	27.8	721	6	BY750607	BY750607 BY750607
27	343.2	27.4	551	4	BM254191	BM254191 515552.MA
28	321.2	25.6	660	6	BY722497	BY722497 BY722497
29	317.6	25.3	643	4	BM684087	BM684087 UT-E-E31-
30	314.4	25.1	701	6	BY734317	BY734317 BY734317
31	302.2	24.1	434	7	NS9044	NS9044 yv60b12.r1
32	299.6	23.9	375	5	BU664580	BU664580 c1119p02.r1
33	294.2	23.5	499	2	BF603905	BF603905 269531.MA
34	288.2	23.0	488	2	BF602079	BF602079 267224.MA
35	280.4	22.4	550	9	CG573938	CG573938 O5T206520
36	273.6	21.8	535	5	BX529358	BX529358 BX529358
37	271.2	21.6	488	4	BG944752	BG944752 ax54d09.x
38	270.4	21.6	416	7	R10548	R10548 yf31e01.r1
39	257.8	20.6	601	5	BU661774	BU661774 c1176g01.z
40	255.8	20.4	522	9	CG634208	CG634208 O5T354881
41	253.6	20.2	343	7	T84327	T84327 yd37a05.r1
42	250	19.9	504	9	CG669296	CG669296 O5T466105
43	243	19.4	515	2	BF191605	BF191605 239237.MA
44	240.8	19.2	474	6	BY563146	BY563146 BY563146
45	240.8	19.2	476	6	BY559962	BY559962 BY559962

ALIGNMENTS

RESULT 1
BX406635
LOCUS
DEFINITION BX406635 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
ACCESSION CS0DM013YA17.5-PRIME, mRNA sequence.
VERSION BX406635
KEYWORDS BX406635.2 GI:46952745
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS Li, M.B., Gruber, C., Jesssee, J., and Polayres, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 15, 2003 this sequence version replaced gi:30762708.
Contact: Genoscope

Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 854.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0AM013AA09QP1c=854.r.

FEATURES

source
1..964
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM013YA17"
/issue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: Liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-Oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN

Query Match 65.7%; Score 824.4; DB 5; Length 964;
Best Local Similarity 93.0%; Pred. No. 6.5e-216;

Matches	857; Conservative	16; Mismatches	47; Indels	2; Gaps	1;
QY	1	ATGAGCTTAAGTACCGGCGTCTGTCGCGCGTGCCTGCTCCCTGGGGCCCTTAACACTG	60		
Db	43	ATGAGCTTAAGTACCGGCGTCTGTCGCGCGTGCCTGCTCCCTGGGGCCCTTAACACTG	102		
QY	61	GAAGAGCTCTCATTTCTCTCTTTCTTAATTTTTTAAACCACTATGACGCTTCTTAGAGAT	120		
Db	103	GAAGAGCTCTCATTTCTCTCTTTCTTAATTTTTTAAACCACTATGACGCTTCTTAGAGAT	162		
QY	121	CAAAAGGGGCTCGTGGCATCTATCAAGTGGGCAAGATCTACCGTATGAGCGCCATT	180		
Db	163	CAAAAGGGGCTCGTGGCATCTATCAAGTGGGCAAGATCTACCGTATGAGCGCCATT	222		
QY	181	GCGTTGGGCTTCTCACTCTGAGTTCCGAGACACAGCTGGAGAGTGGGCTTCAAC	240		
Db	223	GCGTTGGGCTTCTCACTCTGAGTTCCGAGACACAGCTGGAGAGTGGGCTTCAAC	282		
QY	241	CTCTTCATGCTGGCGCTGCTGTCGAGTGGGCAATCTCTGTCGAGCGCTTCTGAGCCAG	300		
Db	283	CTCTTCATGCTGGCGCTGCTGTCGAGTGGGCAATCTCTGTCGAGCGCTTCTGAGCCAG	342		
QY	301	TTCCCTTCTGGGAAAGTGTCTATCACTGTTCAAGTATTCGGCTGGGCCACCATAGAGCT	360		
Db	343	TTCCCTTCTGGGAAAGTGTGTATCACTGTTCAAGTATTCGGCTGGGCCACCATAGAGCT	402		
QY	361	TTGTGGGCTGCTATCTCACTGAGTGTCTTTGGGAAAGTCACTTGGCGCAGTTGGT	420		
Db	403	ATGTGGGCTGCTATCTCACTGAGTGTCTTTGGGAAAGTCACTTGGCGCAGTTGGT	462		
QY	421	GTGATGCTGCTGGTGGAGAGTGAAGCTTTAGGCAACCTGAGAGATGTCATAGTAATATC	480		
Db	463	GTGATGCTGCTGGTGGAGAGTGAAGCTTTAGGCAACCTGAGAGATGTCATAGTAATATC	522		
QY	481	TTCAACACAGACTACCAATGATGATGACATCTACGTTTCGAGCCCTATTTTGGG	540		
Db	523	TTCAACACAGACTACCAATGATGATGACATCTACGTTTCGAGCCCTATTTTGGG	582		
QY	541	CTGCTGCTGGGCTGGTGGCTGCAAGGCTTCAACCGAGGGAACGAGATTAAGATCAG	600		
Db	583	CTGCTGCTGGGCTGGTGGCTGCAAGGCTTCAACCGAGGGAACGAGATTAAGATCAG	642		
QY	601	ACAGCAACGATACCACTTTTGTGTCATGTCGAGCGCCCTCTTCTTGTGAGATTTCTG	660		
Db	643	ACAGCAACGATACCACTTTTGTGTCATGTCGAGCGCCCTCTTCTTGTGAGATTTCTG	702		
QY	661	CCAAAGTTCAACTCTGCTGCTGAGAAATCCAAATCCAAAGAAAGATCCGCTGTTCAAC	720		
Db	703	CCAAAGTTCAACTCTGCTGCTGAGAAATCCAAATCCAAAGAAAGATCCGCTGTTCAAC	762		
QY	721	ACCTATATGCTGAGAGTCAAGTGGTGGAGCAATCTCAAGGTCATCTTGGCTCAC	780		
Db	763	ACCTATATGCTGAGAGTCAAGTGGTGGAGCAATCTCAAGGTCATCTTGGCTCAC	822		
QY	781	CCCCAAGGAAGATCAGCAAGTATATGTCACAGTGGGCTGTTGGCAGAGGCGCGGT	840		
Db	823	CCCCAAGGAAGATCAGCAAGTATATGTCACAGTGGGCTGTTGGCAGAGGCGCGGT	882		
QY	841	GTGGGTAACCTCG--TGTCACTGATCCCTTCTCCGTGGCTTGGCATGCTGGTCTTGG	898		
Db	883	GTGGGTAACCTCGCTGTGTCATCTKAKCCCTTCTCGGGGCTTGSAMSGTGCTCGGCTTC	942		
QY	899	TGGCTGGGCTGATCTCCGTGG	920		
Db	943	TGGCTGGGCTGATCTCAHCG	964		

RESULT 2
 BU664461
 LOCUS
 DEFINITION
 ACCESSION

QY	1	ATGAGCTTAAGTACCGGCGTCTGTCGCGCGTGCCTGCTCCCTGGGGCCCTTAACACTG	60		
Db	43	ATGAGCTTAAGTACCGGCGTCTGTCGCGCGTGCCTGCTCCCTGGGGCCCTTAACACTG	102		
QY	61	GAAGAGCTCTCATTTCTCTCTTTCTTAATTTTTTAAACCACTATGACGCTTCTTAGAGAT	120		
Db	103	GAAGAGCTCTCATTTCTCTCTTTCTTAATTTTTTAAACCACTATGACGCTTCTTAGAGAT	162		
QY	121	CAAAAGGGGCTCGTGGCATCTATCAAGTGGGCAAGATCTACCGTATGAGCGCCATT	180		
Db	163	CAAAAGGGGCTCGTGGCATCTATCAAGTGGGCAAGATCTACCGTATGAGCGCCATT	222		
QY	181	GCGTTGGGCTTCTCACTCTGAGTTCCGAGACACAGCTGGAGAGTGGGCTTCAAC	240		
Db	223	GCGTTGGGCTTCTCACTCTGAGTTCCGAGACACAGCTGGAGAGTGGGCTTCAAC	282		
QY	241	CTCTTCATGCTGGCGCTGCTGTCGAGTGGGCAATCTCTGTCGAGCGCTTCTGAGCCAG	300		
Db	283	CTCTTCATGCTGGCGCTGCTGTCGAGTGGGCAATCTCTGTCGAGCGCTTCTGAGCCAG	342		
QY	301	TTCCCTTCTGGGAAAGTGTCTATCACTGTTCAAGTATTCGGCTGGGCCACCATAGAGCT	360		
Db	343	TTCCCTTCTGGGAAAGTGTGTATCACTGTTCAAGTATTCGGCTGGGCCACCATAGAGCT	402		
QY	361	TTGTGGGCTGCTATCTCACTGAGTGTCTTTGGGAAAGTCACTTGGCGCAGTTGGT	420		
Db	403	ATGTGGGCTGCTATCTCACTGAGTGTCTTTGGGAAAGTCACTTGGCGCAGTTGGT	462		
QY	421	GTGATGCTGCTGGTGGAGAGTGAAGCTTTAGGCAACCTGAGAGATGTCATAGTAATATC	480		
Db	463	GTGATGCTGCTGGTGGAGAGTGAAGCTTTAGGCAACCTGAGAGATGTCATAGTAATATC	522		
QY	481	TTCAACACAGACTACCAATGATGATGACATCTACGTTTCGAGCCCTATTTTGGG	540		
Db	523	TTCAACACAGACTACCAATGATGATGACATCTACGTTTCGAGCCCTATTTTGGG	582		
QY	541	CTGCTGCTGGGCTGGTGGCTGCAAGGCTTCAACCGAGGGAACGAGATTAAGATCAG	600		
Db	583	CTGCTGCTGGGCTGGTGGCTGCAAGGCTTCAACCGAGGGAACGAGATTAAGATCAG	642		
QY	601	ACAGCAACGATACCACTTTTGTGTCATGTCGAGCGCCCTCTTCTTGTGAGATTTCTG	660		
Db	643	ACAGCAACGATACCACTTTTGTGTCATGTCGAGCGCCCTCTTCTTGTGAGATTTCTG	702		
QY	661	CCAAAGTTCAACTCTGCTGCTGAGAAATCCAAATCCAAAGAAAGATCCGCTGTTCAAC	720		
Db	703	CCAAAGTTCAACTCTGCTGCTGAGAAATCCAAATCCAAAGAAAGATCCGCTGTTCAAC	762		
QY	721	ACCTATATGCTGAGAGTCAAGTGGTGGAGCAATCTCAAGGTCATCTTGGCTCAC	780		
Db	763	ACCTATATGCTGAGAGTCAAGTGGTGGAGCAATCTCAAGGTCATCTTGGCTCAC	822		
QY	781	CCCCAAGGAAGATCAGCAAGTATATGTCACAGTGGGCTGTTGGCAGAGGCGCGGT	840		
Db	823	CCCCAAGGAAGATCAGCAAGTATATGTCACAGTGGGCTGTTGGCAGAGGCGCGGT	882		
QY	841	GTGGGTAACCTCG--TGTCACTGATCCCTTCTCCGTGGCTTGGCATGCTGGTCTTGG	898		
Db	883	GTGGGTAACCTCGCTGTGTCATCTKAKCCCTTCTCGGGGCTTGSAMSGTGCTCGGCTTC	942		
QY	899	TGGCTGGGCTGATCTCCGTGG	920		
Db	943	TGGCTGGGCTGATCTCAHCG	964		

RESULT 2
 BU664461
 LOCUS
 DEFINITION
 ACCESSION

685 bp mRNA linear EST 30-SEP-2002
 c117b02.z1 Hembase; Erythroid Precursor Cells (LDB:cl library)
 Homo sapiens CDNA clone c117b02 5', mRNA sequence.

Oy	186	GGGCTTCCTCACTCAGATTTCCGGAGACACAGCTGGACAGTGGCTTCAACCTTT	245
Db	213	GGGCTTCCTCACTCAGATTTCCGGAGACACAGCTGGACAGTGGCTTCAACCTTT	272
Oy	246	CATGCTGGCGCTTGGTGTGCACTGGGGCAATCTCTGTGACGGCTTCTGAGCCAGTTCC	305
Db	273	CATGCTGGCGCTTGGTGTGCACTGGGGCAATCTCTGTGACGGCTTCTGAGCCAGTTCC	332
Oy	306	TTCTGGGAAAGGTGTCATCACCTGTTCAGTATTCGGCTGGGCCACCATAGAGCTTTGTC	365
Db	333	TTCTGGGAAAGGTGTCATCACCTGTTCAGTATTCGGCTGGGCCACCATAGAGCTTATGTC	392
Oy	366	GGTGCTGATCTCAGTGGATGCTGTCTTGGGGAAAGTCACTTGGCGCAGTTGGTGTGAT	425
Db	393	GGTGCTGATCTCAGCGGGGTGCTGTCTTGGGGAAAGTCACTTGGCGCAGTTGGTGTGAT	452
Oy	426	GGTGCTGGTGGAGGTGACAGCTTTTAGGCCAAGCTTAGAGATGTGTCATAGTATATCTTCA	485
Db	453	GGTGTGTGTGGAGGTGACAGCTTTTAGGCCAAGCTTAGAGATGTGTCATAGTATATCTTCA	512
Oy	486	CACAGACTACCAATGAAACATGATGCAACATCTACAGTGTCCGAGCCTATTTTGGGCTGAC	545
Db	513	CACAGACTACCAATGAAACATGATGCAACATCTACAGTGTTCGAGCCTATTTTGGGCTGAC	572
Oy	546	TGTGCTGTGTGCTGCCAAGCCTCTTACCCGAGGGAAACGAGAGATTAAGATTCAGACAC	605
Db	573	TGTGCTGTGTGCTGCCAAGCCTCTTACCCGAGGGAAACGAGAGATTAAGATTCAGAGAC	632
Oy	606	AACGATACCACAGTTTGTGTGTCATGCTGGGGGCGCCTCTTCTGTGGATGTTCT	658
Db	633	AACGATACCACAGTTTGTGTGTCATGCTGGGGGCGCCTCTTCTGTGGATGTTCT	685

RESULT 3	676 bp	mRNA	linear	EST 06-FEB-2003
BX106517				
LOCUS				
DEFINITION	BX106517 Soares fetal liver spleen INFLS Homo sapiens			
	IMAGE998001132 ; IMAGE:128472, mRNA sequence.			

SOURCE ORGANISM	Homo sapiens (human)
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

REFERENCE
AUTHORS
TITLE
JOURNAL
DOI/PID
Copyright: Tm Publisher

COMMENT

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 RZPD: IMAGP98001132.
 RZPDLIB: I.M.A.G.E. cDNA Clone Collection;
 Human Unigeneset - RZPD3 (RZPDLIB No. 972)
[http://www.rzpd.de/CloneCards/cgi-
 bin/showlib.pl.cgi?response=libNo=972](http://www.rzpd.de/CloneCards/cgi-bin/showlib.pl.cgi?response=libNo=972) Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel.: +49 30 32639 101
 Fax: +49 30 32639 111
 www.rzpd.de

FEATURES

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGP998001132 ; IMAGE:128472"

```

```

/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/notes="Organ: Liver and Spleen; Vector: pT733D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with Pac I - oligo(dT) primer
[5' AACTCGACAGATTTAAATTAAGATCTTTTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT733 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fátima Bonafido."

```

Query Match	50.2%	Score 639.8	DB 5	length 676
Best Local Similarity	98.2%	Pred. No. 3e-162		
Matches 637	Conservative 0	Mismatches 12	Indels 0	Gaps 0

1 ATGAGCTCTAAGTACCCGGGCTGTCCGGGCGGTGCTTGGCCCTCTGGGCGCCCTAACACTG 60

ה'תשס"ב

כ"א חשוון ה'תשס"ב

Db 88 GAAGAGCTCTCATTCCTCTTCTATTCTTTTACCCACTATGAGCCTTCCTTAGAGGAT 14

121 CAAAAGGGGCTCGTGGCATCTATCAAGTTGGCCAAGATCTGACCGTGATGGCGGCCATT 18

[illegible]

db 208 GGCTTGGGCTTCCCTCACTCGAGTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCAAC 26

241 CTCTTCAATGCTGGCGCTTGGTGTGCAGTGGGCAATCCGTGACGGCTTCCGTAGCCAC 30

[illegible]

Db 328 TTCCCTTCGGAGGTGTCATCACACTGTTACGATATTCGGCTGGCCACCATGAGTGCT 388

361 TTGTGGTCTGATCTCAGTGGATGCTGTCCTTGGGGAAGTCAACTTGGCGAGTTGGTG 422

388 AAGTCGGTGGTGAATCTTCTGCTTGGGGAAGTCAATCTGGGCGAGTGGTG 44

448 GTGATGTTGCTGTTGAGGTCACAGCTTTAGGACCCGTGAGGATGGTCATCAGTAAATATC 50

481 TTCAACACAGACTTACCATGACATGATGCACATCTACGTGTTGCAGCCTATTTTGGG 54

D5 508 TTCAACAGACTACACACATGAGACCTGAGGCACTCTACGCTGTCGAGCCCTATTTCGG 56

568

601 ACAGCAACGATACCCAGCTTTGTCTGCCATGCTGGGGCCCTCTTCTGT 649

Db 628 AGAGCAACGATACCCAGTTTGTCTGCCATGCTGGGGGCCCTCTTCTTGT 676

RESULT 4

LOCUS	B06559/8	685 bp	mRNA	linear	BS1 30-SEP-20
DEFINITION	c104g11.z1 Hembase; Erythroid Precursor Cells (LCB:c1 library) Hc				

ACCESSION	BU655978
VERSION	BU655978.1
	GI:23368160

REMARKS	SOURCE
201.	Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
Gene Expression in Human Erythroid Precursor Cells
Unpublished (2002)
Contact: Jeffery L. Miller
Laboratory of Chemical Biology
National Institute of Diabetes and Digestive and Kidney Diseases
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
20892, USA
Tel: 301 403 2373
Fax: 301 435 5148
Email: jml@nih.gov
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the
Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
analyses by National Institutes of Health Intramural Sequencing
Center (NISC). More information available at:
http://hembase.nidk.nih.gov
Plates: 04 row: 9 column: 11
Seq primer: 5' lambda-TripLex2 Sequencing Primer.
Location/Qualifiers

FEATURES

source

1..685
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="cl04g11"
/sex="unknown"
/tissue_type="Blood"
/cell_type="Erythroid Precursor Cells"
/mononuclear_cells="Primary Culture of Peripheral Blood Mononuclear Cells"
/dev_stage="Precursor erythroblasts; GPA++"
/lab_host="DH5alpha"
/clone_lib="Hembase; Erythroid Precursor Cells (LCB:cl library)"
/note="Organ: blood; Vector: pTriplEx2; Site:1: SfiI; Site:2: SfiI; A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA library construction kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1 umol/l peptide nucleic acid (PNA) oligos (N-terminal)-biotin-GTC-CAC-CCG-AGG-CTT-G-(C-terminal) and (N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal) . Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC).
Http://www.nisc.nih.gov/."

ORIGIN

Query Match 49.2%; Score 617.2; DB 5; Length 685;
Best Local Similarity 98.0%; Pred. No. 8.9e-159;
Matches 625; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 ATGAGCTTAAGTACCGGGGCTGTCGGCGCTGCTGCTGCGGCGCTTAACACTG 60
DB 48 ATAGGCTTAAGTACCGGGGCTGTCGGCGCTGCTGCTGCGGCGCTTAACACTG 107
QY 61 GAAGAGCTTCAATCTCTCTTATTTTATACCATAGACGCTTCTTAAGAGAT 120
DB 108 GAAGAGCTTCAATCTCTCTTATTTTATACCATAGACGCTTCTTAAGAGAT 167
QY 121 CAAAGGGGCTGTCGATTCATCAAGTTGGCCAAGATCTGACCGTGAATGGCGGCATT 180
DB 168 CAAAGGGGCTGTCGATTCATCAAGTTGGCCAAGATCTGACCGTGAATGGCGGCATT 227
QY 181 GGCTTGAGCTTCTCACTGAGTTCCGAGAGACAGCTGAGAGAGTGGCTTCAAC 240
|||||

Db 228 GGCTTGAGCTTCTCACTGAGTTCCGAGAGACAGCTGAGAGATGGCTTCAAC 287
QY 241 CTCTTCATGCTGGGCGCTTGGTGTGAGTGGCAATCTGCTGAGACGGCTTCCAGCCAG 300
Db 288 CTCTTCATGCTGGGCGCTTGGTGTGAGTGGCAATCTGCTGAGACGGCTTCCAGCCAG 347
QY 301 TTCCCTTCTGGAGAGGTGTGATCAACCTGTTGATTCAGTGGTGGCCATAGATGCT 360
Db 348 TTCCCTTCTGGAGAGGTGTGATCAACCTGTTGATTCAGTGGTGGCCATAGATGCT 407
QY 361 TTGCGGTGTGATTCAGTGTGAGTGGTGTGAGTGGTGGTGGTGGTGGTGGTGGTGG 420
Db 408 GTGTGGTGTGATTCAGTGTGAGTGGTGTGAGTGGTGGTGGTGGTGGTGGTGGTGG 467
QY 421 GTGATGTGTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 480
Db 468 GTGATGTGTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 527
QY 481 TTCAACAGACATCAACATCAATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 528 TTCAACAGACATCAACATCAATGATGATGATGATGATGATGATGATGATGATGATG 587
QY 541 CTGTCTGTGAGCTGTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
Db 588 CTGACTGTGAGCTGTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 647
QY 601 AAGAGACATCAACATCAATGATGATGATGATGATGATGATGATGATGATGATGATG 638
Db 648 AAGAGACATCAACATCAATGATGATGATGATGATGATGATGATGATGATGATGATG 685

RESULT 5
AKO79335
LOCUS
DEFINITION
Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630028A03 product:Rhesus blood group CE and D, full insert sequence.
AKO79335
AKO79335.1 GI:26347922
VERSION
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
Carninci,P. and Hayashizaki,Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
MEDLINE
99279253
PUBMED
10349636
AUTHORS
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE
Normalization and subraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
MEDLINE
20493374
PUBMED
11042159
AUTHORS
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Katsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-Format sequencing pipeline with 384 multicapillary sequencer
JOURNAL
MEDLINE
20530913
PUBMED
11076861
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the

TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS			
PANTOM Consortium.	Functional annotation of a full-length mouse cDNA collection	Nature 409, 685-690 (2001)	5	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	Nature 420, 563-573 (2002)	6 (bases 1 to 1457)			
Adachi, S., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, F., Fukuda, S., Furuno, M., Hanagata, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numasaki, R., Ohno, M., Ohashi, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, Y., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.	Submitted (15-APR-2002) Yoshitake Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsukuba, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues	Please visit our web site for further details.	URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/.	Location/Qualifiers	1. 1457	/organism="Mus musculus"			
/mol_type="mRNA"	/strain="C57BL/6J"	/db_xref="PANTOM_DB:9630028A03"	/db_xref="taxon:10090"	/clone_xref="9630028A03"	/tissue_type="cerebellum"	/clone_lib="RIKEN full-length enriched mouse cDNA library"	/dev_stage="16 days neonate"			
10. 1266	/note="unnamed protein product; Rhesus blood group CE and D (MGD) MGI:1202882, GB NM_011270, evidence: BLASTN, 99%, match=1257)	putative"	/codon_start=1	/protein_id="BAC37610.1"	/db_xref="GI:26347923"	/translation="MGSKPRPSACCPALAEIQTAPSLFCPEFIHPDRAVDHPRM ASYVLRNLTLMALGEGFTSSSRPRSSMSVATNLFELAGVGCTILLDFLQVQLQ WNKNNLSIIQIATMSTLPVLISAGAVLGKLVGLVTVLMEMMAAGARLFADEKVF KMTKHIIIMHGVHGAFFGLVVAWMLSSLPFRVGENAQTKEVMAKSSSLFPAELGTL FLVIFWPAINSGALLEGTKRNAVENTYATLALVAVTATSMALSHPOKIMVHIIINA PLAVGVGARGCLISSPMIWMVGLIAGLISGACCPACLNHMLONSGIHVTGG LRGILGATVYDCIOTVEPKSSDIWITQYTHIGALSFAYAMQVGLTGLCLLSVR VWRAPRAKTYFDYDDTFWFEFLVANGP"	1436. 1441	/note="putative"	1457	/note="putative"
48.0%; Score 602.2; DB 3; Length 1457;	69.7%; Pred. No. 1.6e-154;									

	Matches	884: Conservative	0: Mismatches	358: Indels	27: Gaps	4
QY	1	ATGAGCTTAAGTACCGCGGATGTGTCTCGGCGCTGCTGCTGCGGCTTAACACTG	60			
DB	10	ATGGGCTCTAAGTACCCCAAGGTCCTCCGCTGCTGCTGCTGCGGCTTAAGACTA	69			
QY	61	GAAGCAGTCTCATTTCTCTCTTCTATTTTTCATGCCATAGACGCTTCTTAAGAT	120			
DB	70	CAGACAGTTTATGCTCTCTCTCTGTTTTTCATCCCCACGACACAGCCAGGTGAT	129			
QY	121	CAAAAGGGGCTCGTGGCATCTCATCAAGTTGGGCCAAGATCTGAACCGTATGGCGGCAAT	180			
DB	130	CACA---GGTTCAATGGCGAGCTATCAAGTCTCTCGGAATTTGACCTCATGGCAGCTTG	186			
QY	181	GAGTTGGGCTTCCTCACCTCGAGTTTCCGAGACACAGCTGGAGCAGTGTGGCTTTCAAC	240			
DB	187	GAGCTTCGGCTTCCTGTCTCTCGCTTTTGGAGACACAGCTGGAGCAGTGTGGCTTTCAAC	246			
QY	241	CTCTTCATGCTGGCGCTTGTGTGCAGTGGGCAATCTGCTGAGACGCTTCTGAGCCAG	300			
DB	247	CTCTTCATGTTGGCGCTTCGGGGGTGAGGGAAACAATCTTGAGCAATTTCTTGGGCCAG	306			
QY	301	TTCCCTTGTGGAAAGGTGATCATCACTGATTAAGTATTCGGCTGGCCACCATGATGCT	360			
DB	307	GTCCTCCATGGAAACATATCAATTTGTCTCCAGATTCAGATTAAGTATGACAC	366			
QY	361	TTGTTCGAGTCAATCTCAGTGGATGCTGCTTGGGGAAAGTCAATTGGCGCATTTGGTG	420			
DB	367	TTACTGTGCTGATCTCAGCGGGCGCTGTCTCTGGGGAAAGTCAACTGTGTGACGTGACC	426			
QY	421	GTCATGTGCTGTGGAGGTGACAGCTTTAGGCAACTGAGAGTGTCACTAGTAATATC	480			
DB	427	GTCATGTGCTGTAGAGGAGCAATGAGCTTTGTGTGCATCAATTTGCCGACGAGAAAGTTC	486			
QY	481	TTCAACACAGACTACATATGAACATGTGACATCTACGTGTTGGCAGCCATATTTGGG	540			
DB	487	TTCAAAATGACAGAACATATCATCTATGTATGACAGGGCAGTGTGTGGGCTCATATTTGGG	546			
QY	541	CTGTCTGTGGCGTGGTGCCTGCCAAAGGCTCTACCCGAGGGAAAGGAAAGATCAAGATCAG	600			
DB	547	CTAATGTGGCTTGGTGGCTTTCCAGATCTCTGCCAAGAGAGTGGTGTGAAACGCCACAG	606			
QY	601	ACA-----GCACGATACCAAGTTGTCTGCCATGCTGGCGGCTCTTTC	645			
DB	607	ACAGAGAAAGTTCAAAATGGCTACGAGCTCCAGTCTGTTTGCCATGCTGGGCACTCTTTC	666			
QY	646	TTGTGAGATTTCTGGCCCAAGTTTCACTCTGCTGTGCTGAGAAATGTCATTCGAAGAAG	705			
DB	667	TTGTGAGATTTCTGGCCAGCTATCAACTCTCTCTCTGGAAGGGACAAAGAAAG---G	723			
QY	706	AATGCGGTTCAACACCTTATGCTGTAGAGTACAGTACAGTGGTGAAGCCATCTCAGGG	765			
DB	724	AATGCTGTGTTCAACACCTTATGAGCCCTTGGGAGTGAAGCCAGTACAGCCACTTCCATG	783			
QY	766	TCAATCTTTGGCTCAACCCCAAGGAGATCAAGACACTTATGTGACAGTGGGTGTTG	825			
DB	784	TCAGCCCTGAAGTCAACCTTCMAAGGAGATCAACAGTTCACATCCACAAATGCAAGTGTG	843			
QY	826	GCAGAGGCGTGGCGTGTGGTAACTCGTGTCACTGATTCCTTCTCGGTGCTTGGCATG	885			
DB	844	GCAGGGGCGTGGCGGTGGGCGCCCGGGGTGCTGATTTCTTCACTTTGGAATTTTCCATG	903			
QY	886	GTCGCTGGGCTTTGTGGCTGGGCTGATCTACCGTGGGGGAGGACATCTGCGGGGGTGT	945			
DB	904	GTCCTGGGCTCATATGCTGGGTGATTTCTTCACTGTGGGAGCCGAACTGTCCAGGGGCTGT	963			
QY	946	TGTAAACGAGTCTGGGGATTCGCCACAGCTCACTCATGAGGCTACAACTTCAAGCTTGTG	1005			
DB	964	TTGAACCAATGCTGCAAGACTCCAGTGG-----GATTCGACTACACTTGGGCTTGGCG	1011			
QY	1006	GCTCTGCTGGAGAGATCATTTACATTTGTGCTGCTGTGCTTGAATACCGTGGAGCGGCG	1061			
DB	1018	GGTCTGCTGGAGAGCACTTACTTACTTCTGCTTCAAGATGTGACAGAGCCCAAGTCTTCG	1071			

QY 1066 AATGGCATATGGCTTCCAGGCTCTCAGCATTTGGGAACTGAGCTTGGCCATGCG 1125
 DB 1078 GATCTCTGATATCATCCAGACGGTCACTCATCTGGGCTCTGAGCTTGGCTGGCG 1137
 QY 1126 ATAGCTCTACGCTGTGTCTCTGACAGGTTGCTCTTAAATCTTAAATATGAAAGCA 1185
 DB 1138 ATGGGTATGATGATCTGAGCTCTCTCAGGTTGTCTCTTAAAGTGCAGAGTGGAGGCT 1197
 QY 1186 CCTCATGAGGCTAAATATTTTATGATGACCAAGTTTCTGGAAGTTCTCTCATTTGGCTGT 1245
 DB 1198 CCCCCTGGCGCCAAAGATTTTATGATATGACATTTCTGGAGTTCCACACTTGGCGGT 1257
 QY 1246 GGATTTTAA 1254
 DB 1258 GGATTTTAA 1266

RESULT 6
 BU661907
 LOCUS BU661907 706 bp mRNA linear EST 30-SEP-2002
 DEFINITION c178e10.21 Hembase; Erythroid Precursor Cells (LCB:c1 library) Homo
 sapiens cDNA clone c178e10 5', mRNA sequence.
 BU661907
 ACCESSION BU661907.1 GI:23374089
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Gubin, A.N., Lee, Y.T., Bouffard, G.G. and Miller, J.L.
 Gene Expression in Human Erythroid Precursor Cells
 Unpublished (2002)
 Contact: Jeffrey L. Miller
 Laboratory of Chemical Biology
 National Institute of Diabetes and Digestive and Kidney Diseases
 Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
 20892, USA
 Tel: 301 402 2373
 Fax: 301 435 5148
 Email: jlm7f@nih.gov
 The 'c1' library was constructed by Alexander Gubin, Ph.D. in the
 Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
 analyses by National Institutes of Health Intramural Sequencing
 Center (NISC). More information available at:
 http://hembase.nidk.nih.gov
 Plate: 78 row: e column: 10
 Seq primer: 5' lambda-TripLex2 Sequencing Primer.
 Location/Qualifiers
 1..706
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="c178e10"
 /sex="unknown"
 /issue_type="blood"
 /cell_type="Erythroid Precursor Cells"
 /cell_line="Primary Culture of Peripheral Blood
 Mononuclear Cells"
 /dev_stage="Precursor erythroblasts; GPA++"
 /lab_host="DH5alpha"
 /clone_lib="Hembase; Erythroid Precursor Cells (LCB:c1
 library)"
 /note="Organ: blood; Vector: pTriplex2; Site: 1: SfiI;
 Site: 2: SfiI; A complementary DNA (cDNA) library from
 human erythroid precursor cells was constructed using
 SMART PCR (polymerase chain reaction) cDNA library
 Construction Kit (Clontech, Palo Alto, CA) according to
 the manufacturer's directions, but with slight
 modifications. Briefly, reverse transcription was
 performed in the presence of 1 umol/L peptide nucleic acid
 (PNA) oligos

(N-terminal)-biotin-GTC-CAC-CCG-AMG-CTT-G-(C-terminal) and
 (N-terminal)-biotin-C(T/C)T-GAA-GTT-CTC-AGG-A-(C-terminal)
 . Synthesized cDNA was digested with SfiI and
 size-selected on a 1% agarose gel (>800bp). Large-scale
 sequencing of the library was performed by the NIH
 Intramural Sequencing Center (NISC).
 Http://www.nisc.nih.gov/."

ORIGIN

Query Match 47.6%; Score 597.4; DB 5; Length 706;
 Best Local Similarity 97.4%; Pred. No. 2,6e-153;
 Matches 607; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATAGGCTCTAAGTACCCGGGCTGTGTCCGGCTGCTGCGCCCTTGGGCGCTTAACATG 60
 DB 84 ATAGGCTCTAAGTACCCGGGCTGTGTCCGGCTGCTGCGCCCTTGGGCGCTTAACATG 143
 QY 61 GAAGCAGCTCTCATTTCTCTTCTTATTTTATACCACTAAGAGCTTCTTAAGAGAT 120
 DB 144 GAAGCAGCTCTCATTTCTCTTCTTATTTTATACCACTAAGAGCTTCTTAAGAGAT 203
 QY 121 CAAAGGGGCTGTGGGATCTTATCAAGTTGGCCAAAGATCTGACCCGTATGGCGCAT 180
 DB 204 CAAAGGGGCTGTGGGATCTTATCAAGTTGGCCAAAGATCTGACCCGTATGGCGCAT 263
 QY 181 GGCTTGGGCTTCTCACTCCAGTTTCCGAGACACACACTGGAGAGTGTGGCTTCAAC 240
 DB 264 GGCTTGGGCTTCTCACTCCAGTTTCCGAGACACACACTGGAGAGTGTGGCTTCAAC 323
 QY 241 CTCTTCATGCTGGCCCTTGTGTGAGTGGGCAATCTGCTGAGACGGCTTCTAGAGCAG 300
 DB 324 CTCTTCATGCTGGCCCTTGTGTGAGTGGGCAATCTGCTGAGACGGCTTCTAGAGCAG 383
 QY 301 TTCCCTTCTGGGAAGTGTGTATCACTGTTCAAGTATTTGGCTGGCCACATGAGTCT 360
 DB 384 TTCCCTTCTGGGAAGTGTGTATCACTGTTCAAGTATTTGGCTGGCCACATGAGTCT 443
 QY 361 TTGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
 DB 444 ATGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 503
 QY 421 GTGATGT 480
 DB 504 GTGATGT 563
 QY 481 TTCAACACAGACTACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 540
 DB 564 TTCAACACAGACTACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 623
 QY 541 CTGTCTGTGCTGT 600
 DB 624 CTGACTGT 683
 QY 601 ACAGCAACGATACCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 623
 DB 684 AGAGCAACGATACCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 706

FEATURES
 source
 1..706
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="c178e10"
 /sex="unknown"
 /issue_type="blood"
 /cell_type="Erythroid Precursor Cells"
 /cell_line="Primary Culture of Peripheral Blood
 Mononuclear Cells"
 /dev_stage="Precursor erythroblasts; GPA++"
 /lab_host="DH5alpha"
 /clone_lib="Hembase; Erythroid Precursor Cells (LCB:c1
 library)"
 /note="Organ: blood; Vector: pTriplex2; Site: 1: SfiI;
 Site: 2: SfiI; A complementary DNA (cDNA) library from
 human erythroid precursor cells was constructed using
 SMART PCR (polymerase chain reaction) cDNA library
 Construction Kit (Clontech, Palo Alto, CA) according to
 the manufacturer's directions, but with slight
 modifications. Briefly, reverse transcription was
 performed in the presence of 1 umol/L peptide nucleic acid
 (PNA) oligos

RESULT 7
 BX406634/c
 LOCUS BX406634/c 878 bp mRNA linear EST 03-MAY-2004
 DEFINITION BX406634 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
 CS0DM013YA17 3-PRIME, mRNA sequence.
 ACCESSION BX406634
 VERSION BX406634.2 GI:46951844
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 IJ, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization

ORIGIN

sequencing of the library was performed by the NIH
Intramural Sequencing Center (NISC;
Http://www.nisc.nih.gov/).

Query Match 44.7%; Score 560.4; DB 5; Length 631;
Best Local Similarity 99.6%; Pred. No. 4.1e-143;
Matches 561; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAGCTTAAGTACCCGCGGTCTGTCCGGGCGTCCCTCTGGGCGCTTAACATG 60
DB 69 ATGAGCTTAAGTACCCGCGGTCTGTCCGGGCGTCCCTCTGGGCGCTTAACATG 128
QY 61 GAAGAGCTCTCAATCTCTCTTTATTTTAACTAATGAGCTTCTTGAAGAGAT 120
DB 129 GAAGAGCTCTCAATCTCTCTTTATTTTAACTAATGAGCTTCTTGAAGAGAT 188
QY 121 CAAAGAGGGGCTGTGGGCAATCCATAGTGGCCCAAGATCTGACCTGATGGCGCAT 180
DB 189 CAAAGAGGGGCTGTGGGCAATCCATAGTGGCCCAAGATCTGACCTGATGGCGCAT 248
QY 181 GGGTTGGGCTTCTCACTCGAGTTCCGAGACACAGCTGAGAGAGTGGCGCTTGAAC 240
DB 249 GGGTTGGGCTTCTCACTCGAGTTCCGAGACACAGCTGAGAGAGTGGCGCTTGAAC 308
QY 241 CTCTTCATGCTGCGCTTGTGTGCAGTGGCAATCTGCTGGAAGGCTTCTGAGCGAG 300
DB 309 CTCTTCATGCTGCGCTTGTGTGCAGTGGCAATCTGCTGGAAGGCTTCTGAGCGAG 368
QY 301 TTCCCTTCTGGGAAGGTGTCATCACTGTCAGTATTTGGGCGGCAACATAGAGCT 360
DB 369 TTCCCTTCTGGGAAGGTGTCATCACTGTCAGTATTTGGGCGGCAACATAGAGCT 428
QY 361 TTGTGGGCTGATCTCAGTGGATCTGTCTTGGGGAAGGTCACTTGGCGAGTTGGT 420
DB 429 TTGTGGGCTGATCTCAGTGGATCTGTCTTGGGGAAGGTCACTTGGCGAGTTGGT 488
QY 421 GTATGAGGCTGTGGGAGGTGACAGCTTTAGGCACTGAGAGATGTCATCAGTATATC 480
DB 489 GTATGAGGCTGTGGGAGGTGACAGCTTTAGGCACTGAGAGATGTCATCAGTATATC 548
QY 481 TTCAACAAGATCACTCAATGATGATGACATCTACGTTGTCGAGCTTATTTGG 540
DB 549 TTCAACAAGATCACTCAATGATGATGACATCTACGTTGTCGAGCTTATTTGG 608
QY 541 CTGTCTGTGGCTGTGGTGGCTGCC 563
DB 609 CTGTCTGTGGCTGTGGTGGCTGCC 631

```

RESULT 9
BU665450 632 bp mRNA linear EST 30-SHP-2002
LOCUS cl131d04.z1 Hembase: Erythroid Precursor Cells (LCB:cl1 library)
DEFINITION Homo sapiens cDNA clone cl131d04 5', mRNA sequence.
ACCESSION BU665450
VERSION BU665450.1 GI:23377637
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 632)
Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.
Gene Expression in Human Erythroid Precursor Cells
Unpublished (2002)
Contact: Jeffery L. Miller
Laboratory of Chemical Biology
National Institute of Diabetes and Digestive and Kidney Diseases
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
20892, USA
Tel: 301 402 2373
Fax: 301 435 5148

FEATURES

source

Email: jmf7@nih.gov
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the
Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or
analyses by National Institutes of Health Intramural Sequencing
Center (NISC). More information available at:
http://hembase.nidk.nih.gov
Plate: 131 row: d column: 04
Seq primer: 5' Lambda-p10r1ex2 Sequencing Primer.
Location/Qualifiers
1..632
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="cl131d04"
/sex="unknown"
/tissue_type="blood"
/cell_type="Erythroid Precursor Cells"
/cell_line="Primary Culture of Peripheral Blood
Mononuclear Cells"
/dev_stage="Precursor erythroblasts; GPA++"
/lab_host="DMSalpa"
/clone_lib="Hembase; Erythroid Precursor Cells (LCB:cl
library)"
/note="Organ: blood; Vector: pTriplex2; Site 1: SfiI;
Site 2: SfiI; A complementary DNA (cDNA) library from
human erythroid precursor cells was constructed using
SMART PCR (polymerase chain reaction) cDNA library
construction kit (Clontech, Palo Alto, CA) according to
the manufacturer's directions, but with slight
modifications. Briefly, reverse transcription was
performed in the presence of 1 umol/L peptide nucleic acid
(PNA) oligos
(N-terminal)-biotin-GTC-CAC-CCG-AGG-CCT-G-(C-terminal) and
(N-terminal)-biotin-C(T/C)-T-GAA-GTT-CTC-AGG-A-(C-terminal)
. Synthesized cDNA was digested with SfiI and
size-selected on a 1% agarose gel (>800bp). Large-scale
sequencing of the library was performed by the NIH
Intramural Sequencing Center (NISC;
Http://www.nisc.nih.gov/).

ORIGIN

Query Match 44.6%; Score 559; DB 5; Length 632;
Best Local Similarity 98.3%; Pred. No. 1e-142;
Matches 565; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

QY 1 ATGAGCTTAAGTACCCGCGGTCTGTCCGGGCGTCCCTCTGGGCGCTTAACATG 60
DB 58 ATGAGCTTAAGTACCCGCGGTCTGTCCGGGCGTCCCTCTGGGCGCTTAACATG 117
QY 61 GAAGAGCTCTCAATCTCTCTTTATTTTAACTAATGAGCTTCTTGAAGAGAT 120
DB 118 GAAGAGCTCTCAATCTCTCTTTATTTTAACTAATGAGCTTCTTGAAGAGAT 177
QY 121 CAAAGAGGGGCTGTGGGCAATCCATAGTGGCCCAAGATCTGACCTGATGGCGCAT 180
DB 178 CAAAGAGGGGCTGTGGGCAATCCATAGTGGCCCAAGATCTGACCTGATGGCGCAT 237
QY 181 GGGTTGGGCTTCTCACTCGAGTTCCGAGACACAGCTGAGAGAGTGGCGCTTGAAC 240
DB 238 GGGTTGGGCTTCTCACTCGAGTTCCGAGACACAGCTGAGAGAGTGGCGCTTGAAC 297
QY 241 CTCTTCATGCTGCGCTTGTGTGCAGTGGCAATCTGCTGGAAGGCTTCTGAGCGAG 300
DB 298 CTCTTCATGCTGCGCTTGTGTGCAGTGGCAATCTGCTGGAAGGCTTCTGAGCGAG 357
QY 301 TTCCCTTCTGGGAAGGTGTCATCACTGTCAGTATTTGGGCGGCAACATAGAGCT 360
DB 358 TTCCCTTCTGGGAAGGTGTCATCACTGTCAGTATTTGGGCGGCAACATAGAGCT 417
QY 361 TTGTGGGCTGATCTCAGTGGATCTGTCTTGGGGAAGGTCACTTGGCGAGTTGGT 420
DB 418 ACGTGGGCTGATCTCAGTGGGCTGTCTTGGGGAAGGTCACTTGGCGAGTTGGT 477

```


Query Match	43.3%	Score 542.4	DB 5	Length 607
Best Local Similarity	98.0%	Pred. No. 3.8e-118		
Matches 545	Conservative 0	Mismatches 11	Indels 0	Gaps 0
QY	1	ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGCTGCTGCCCTCTGTGGGCCCTAACA	CTG	60
Db	48	ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGCGCTGCTGCCCTCTGTGGGCCCTAACA	CTG	107
QY	61	GAAGCAGCTCTCATTTCTCTCTTTATTTTTTTTAAACCACTATGACGCTTCTTAGAGAT		120
Db	108	GAAGCAGCTCTCATTTCTCTCTTTATTTTTTTTAAACCACTATGACGCTTCTTAGAGAT		167
QY	121	CAAAAGGGGGCTCGGGGATCCTATCAAGTTGGGCCAAGATCTGACCGGTATGGGGCCATT		180
Db	168	CAAAAGGGGGCTCGGGGATCCTATCAAGTTGGGCCAAGATCTGACCGGTATGGGGCCATT		227
QY	181	GGCTTGGGCTTCCCTACCTCGAGTTTCCGAGACACACAGCTGGACAGTGGGCTTCAAC		240
Db	228	GGCTTGGGCTTCCCTACCTCGAGTTTCCGAGACACACAGCTGGACAGTGGGCTTCAAC		287
QY	241	CTCTTCATGCTGCGCTTGGTGTGACAGTGGGCAATCCTGCTGACCGCTTCTGAGCCAG		300
Db	288	CTCTTCATGACAGGCGCTTGGTGTGACAGTGGGCAATCCTGCTGACCGCTTCTGAGCCAG		347
QY	301	TTCCCTCTTGGGAAGGTGATCATACACTGTTCAGTATTCGGCGTGGCCACATAGTCT		360
Db	348	TTCCCTCTTGGGAAGGTGATCATACACTGTTCAGTATTCGGCGTGGCCACATAGTCT		407
QY	361	TTGTCGGTGCATCTCATGTGATGTCTGTCTTGGGGAAAGTCAACTTGGGCGCATTTGGTG		420
Db	408	ATGTCGGTGCATCTCATGTGATGTCTGTCTTGGGGAAAGTCAACTTGGGCGCATTTGGTG		467
QY	421	GTGATGGTCTGTGTGAGGTGACAGCTTTAGGCAACTGAGAGTGTTCATCATTAATATC		480
Db	468	GTGATGGTCTGTGTGAGGTGACAGCTTTAGGCAACTGAGAGTGTTCATCATTAATATC		527
QY	481	TTCAACACACACTACCATATGAATATGACACTTACGTGTTCGACGCTATTTTGGG		540
Db	528	TTCAACACACACTACCATATGAATATGACACTTACGTGTTCGACGCTATTTTGGG		587
QY	541	CTGTCGTGGCGCTGGTGCTT		560
Db	588	CTGACTGTGGCGCTGGTGCTT		607

RESULT 13

AK089642

LOCUS

DEFINITION

AK089642

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

FEATURES	COMMENT	TITLE	AUTHORS
CD5		JOURNAL REFERENCE	<p>SHIBATA, K., ITOH, M., AIZAWA, K., NAGASAKI, S., SASAKI, N., CARINCI, P., KONNO, H., ARIYAMA, J., NISHI, K., KISUMAI, T., TASHIRO, H., ITOH, M., SUMI, N., ISHII, Y., NAKAMURA, S., HAZAMA, M., NISHINE, T., HARADA, A., YAMAMOTO, R., MATSUMOTO, H., SAKAGUCHI, S., IKEGAMI, T., KASHIWAGI, K., FUJIIWAKE, S., INOUE, K., TOGAWA, Y., IZAWA, M., OHARA, E., WATAHIKI, M., YONEDA, Y., ISHIIKAWA, T., OZAWA, K., TANAKA, T., MATSUURA, S., KAWAI, J., OKAZAKI, Y., MURAMATSU, M., INOUE, Y., KITAH, A. and HAYASHIZAKI, Y.</p> <p>RIKEN integrated sequence analysis (RISA) system--364-format sequencing pipeline with 384 multicapillary sequencer</p> <p>Genome Res. 10 (11), 1757-1771 (2000)</p> <p>20530913</p> <p>11076861</p>
		JOURNAL REFERENCE	<p>4</p> <p>THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE II TEAM AND THE FANTOM CONSORTIUM.</p> <p>Functional annotation of a full-length mouse cDNA collection</p> <p>Nature 409, 685-690 (2001)</p>
		JOURNAL REFERENCE	<p>5</p> <p>THE FANTOM CONSORTIUM AND THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE I & II TEAM.</p> <p>Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs</p> <p>Nature 420, 563-573 (2002)</p> <p>6 (bases 1 to 975)</p>
		JOURNAL REFERENCE	<p>ADACHI, J., AIZAWA, K., AKIMURA, T., ARAKAWA, T., BONO, H., CARINCI, P., FUKUDA, S., FURUKAWA, M., HANAGAKI, T., HARA, A., HOSHIZUMI, W., HAYASHIDA, K., HAYATSU, N., HIYAMOTO, K., HIRAOKA, T., HIROZANE, T., HORI, F., IMOTO, K., ISHII, Y., ITOH, M., KAGAWA, I., KASUKAWA, T., KATO, S., KAWAI, J., KOJIMA, Y., KONDO, S., KONNO, H., KOUDE, M., KOYA, S., KUIHARA, C., MATSUYAMA, T., MIYAZAKI, A., MURATA, M., NAKAMURA, M., NISHI, K., NOMURA, K., NUMAZAKI, R., OHNO, M., OHNATO, N., OKAZAKI, S., SATO, R., SATOH, H., SAKAI, C., SAKAI, K., SAKAZUME, N., SANO, H., SAAKI, D., SHIBATA, K., SHINAGAWA, A., SHIRAKI, T., SOGABE, Y., TAGAMI, M., TAGAWA, A., TAKAHASHI, F., TAKAKU-AKIHARA, S., TAKEKI, Y., TANAKA, T., TOMARU, A., TOYA, T., YASUNISHI, A., MURAMATSU, M. and HAYASHIZAKI, Y.</p>
		JOURNAL REFERENCE	<p>DIRECT SUBMISSION</p> <p>Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-rs@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)</p>
		JOURNAL REFERENCE	<p>CNA LIBRARY WAS PREPARED AND SEQUENCED IN MOUSE GENOME ENCycloPEDIA Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.</p> <p>Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.</p> <p>Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.</p> <p>Please visit our web site for further details.</p> <p>URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers</p>
		JOURNAL REFERENCE	<p>1. 975</p> <p>/organism="Mus musculus"</p> <p>/mol_type="mRNA"</p> <p>/script="MOD"</p> <p>/db_xref="FANTOM_DB:F830006J05"</p> <p>/db_xref="taxon:10090"</p> <p>/clone="F830006J05"</p> <p>/cissue_type="activated spleen"</p> <p>/clone_lib="RIKEN full-length enriched mouse cDNA library"</p> <p>19. 5975</p> <p>/note="unnamed protein product; Rhesus blood group CE and D (MGD GJI:1202882, GB NM_011370, evidence: BLASTN, 99%, match=1257)</p> <p>putative"</p> <p>/codon_start=1</p> <p>/protein_id="BAC40937.1"</p>

/db_xref="GI:26354621"
/translation="MGSKYPRLRCCLPLMALELOTAFSLFCFPIPHDTQVDRHFM
ASYOVLRLTMAALGFGFLSSPRHSSVAFLNMLAGVOTIILDFLGVLQ
NKKINLSIOIATMSLTPVLSAGAVKXNVYVLTVMLEAFAPLIFADKVF
KMTHEHLYNHGFGAVYRSLTYAMWLSLSPRRVENVQTEKVOATSSLSFLMIGTL
PLMTWPAIINSLBLSIKRNVAVNTYALAVSAVTAIAMSALSHPOGKINMVIHNA
VLAAGVAVGAPGLISPWISWVLIIGLISIWGAKCPRALC"

ORIGIN

Query Match 39.8%; Score 498.6; DB 3; Length 975;
Best Local Similarity 72.4%; Pred. No. 5,4e-126;

Matches 696; Conservative 0; Mismatches 244; Indels 21; Gaps 3;

1 ATGAGCTTAAGTACCCGCGGTCTGTCCGCGCTGCTCCCTCGGAGCCTTAACATG 60
19 ATGGGCTTAAGTACCCGCGGTCTGTCCGCGCTGCTCCCTCGGAGCCTTAACATG 78
61 GAAGGAGCTTCT 120
79 CAGACAGCTTTTAACT 138
121 CAAAAGGGGCTGTGGCATCTATCAAGTTGGCCAAAGTCTGACCGTATGGCGGCATT 180
139 CACA---GGTTCATGGCAGGATCAAGTCTCCGGAATTGACCTCATATGGCAGCCTTG 195
181 GGCTTGGGCTTCTCACTCGAGTTCCGAGAACAGCTGGAGCAAGTGGGCTTCAAC 240
196 GGCTTGGGCTTCT 255
241 CTCTTCATGTGGGCTGT 300
256 CTCTTCATGTGGGCTGT 315
301 TTCCCTTCTGGGAAAGGT 360
316 GTCTTCCAAATGGAACAAGATCAATCTGTCCAGATCCAGATAGCTACATAGAGACC 375
361 TTGTGGGCTGT 420
376 TTACTGT 435
421 GT 480
436 GT 495
481 TTGCAACAACACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
496 TTCAAAATGACAGAAACATCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 555
541 CTGT 600
556 CTAACTGT 615
601 ACA-----GCAACGATACCCAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 645
616 ACAGAGAGGTTCAAAATGCTAGAGCTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 675
646 TTGT 705
676 TTGT 732
706 AATGCGGTGTCAACACTATGT 765
733 AATGCTGTGTCAACACTATGT 792
766 TATCTTGT 825
793 TAGCGCTGT 852
826 GAGAGAGGCTGT 885
853 GAGAGAGGCTGT 912

QY 886 GTGTGGGCTTGT 945
DB 913 GTGTGGGCTTGT 972
QY 946 T 946
DB 973 T 973

RESULT 14
CD627633/c 458 bp mRNA linear EST 12-JUN-2004
LOCUS 56029924H1 FLP Homo sapiens cDNA, mRNA sequence.
DEFINITION CD627633
ACCESSION CD627633
VERSION CD627633.1 GI:40275899
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 458)
Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
Genomics 94 (1), 205-210 (2004)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.

FEATURES
source
1..458
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_11b="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN

Query Match 34.2%; Score 429.2; DB 6; Length 458;
Best Local Similarity 98.2%; Pred. No. 6.1e-107;

Matches 434; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

1 ATGAGCTTAAGTACCCGCGGTCTGTCCGCGCTGCTCCCTCGGAGCCTTAACATG 60
442 ATGAGCTTAAGTACCCGCGGTCTGTCCGCGCTGCTCCCTCGGAGCCTTAACATG 383
61 GAAGGAGCTTCT 120
382 GAAGGAGCTTCT 323
121 CAAAAGGGGCTGTGGCATCTATCAAGTTGGCCAAAGTCTGACCGTGTGGGCGCAT 180
322 CAAAAGGGGCTGTGGCATCTATCAAGTTGGCCAAAGTCTGACCGTGTGGGCGCAT 263
181 GGCTTGGGCTTCT 240
262 GGCTTGGGCTTCT 203
241 CTCTTCATGTGGGCTGT 300
202 CTCTTCATGTGGGCTGT 143
301 TTCCCTTCTGGGAAAGGT 360
142 TTCCCTTCTGGGAAAGGT 83
361 TTGT 420
82 TTGT 23
421 GT 442

This Page Blank (uspto)